

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:06:51 ; Search time 1507.88 Seconds  
(without alignments)  
7418.598 Million cell updates/sec

Title: US-09-671-050-9

Perfect score: 1041

Sequence: 1 atggaaaagtgtgaaaatt.....aggcttcgcgtcaaaagt 1041

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size : 20

Total number of hits satisfying chosen parameters: 6380

inimum DB seq length: 0  
aximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : EST:\*

```
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estcom:*
5: em_estcpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rnd:*
20: em_gss_vrt:*
21: em_gss_other:*
```

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	198	19.0	295	10	AA626859
2	47	4.5	756	11	BG390099
3	41	3.9	433	11	BF869052
4	39	3.7	328	13	AQ230677
5	37	3.6	171	11	B1040145
6	37	3.6	368	13	AQ030066
7	37	3.6	494	13	B32077
8	37	3.6	700	13	AG014845
9	36	3.5	207	10	BE150038
10	36	3.5	366	11	H95588
11	36	3.5	408	10	AA678313
12	36	3.5	500	13	AQ769170

A0668860	HS_5431_A
BG698963	602703384
BE149958	RC0-HT025
AQ617734	HS_5161_B
BF964650	60267728
AQ321133	RPC111-11
BF878219	CM2-ET012
AQ346239	RPC111-11
AM864431	PM4-SN001
AQ093501	HS_3018_B
AW958514	EST370584
BE296260	601174838
A1051456	OW29b12.S
AW339622	he15c05.x
A1244758	q197g12.x
R98160	yq66h03.r1
AQ142429	HS_2244_A
AQ455559	HS_5157_B
M78087	EST01677_Su
AA457230	as91c02.s
BE146457	QV0-HT021
BF802026	CM1-C1009
AA984191	am82b06.s
BF982842	602305602
AQ595304	HS_5448_A
BG575508	602598919
BG912455	602806793
BG151645	na168c04
AQ095034	HS_3028_A
AG002404	Homo sapi
AG002403	Homo sapi
AL048065	DKFp586G
AG002385	Homo sapi
AQ742378	HS_5382_B
AA640211	na95h08.S
AQ341086	RPC111-11
A1814697	WK66d08.x
AL120748	DKFp2762M
AL134562	DKFp547B
BG701158	602680728
BG400082	60242158
BE547155	601072573
BF966168	602286457
N74894	za88c04.s1
AA503246	ne45g08.s
AA302327	EST15447
A1567388	tp67d03.x
AA525875	ni157e09.s
BT007270	MR1-RT007
AA232416	zr23d03.r
A1287787	qv68f05.x
B43053	HS-1057-A1-
BE089717	QV0-BT070
A1254242	qt90f03.x
AA831906	oe21f04.s
A1204643	q044a04.x
AA826078	oe24b03.s
AA666327	ac43c08.s
AW168900	x114h10.x
AW675053	ba60g11.y
N66948	za84e10.s1
A1927393	wo75b10.x
AA626446	ab49h03.r
A1475902	tc95f02.x
AA157881	zo69g09.s
A1753334	cr08f09.x
AW972786	EST384881
AQ756485	HS_5354_B
N70325	za59h12.s1
AA601240	noi14g03.s
AQ190209	HS_3216_A
B70510	CIT-HSP-205

c 86	30	2.9	542	10	AI344320	AI344320 tc03e11.x	159	28	2.7	279	10	AA302660	AA302660 EST10167
c 87	30	2.9	557	10	AV717771	AV717771 AV717771	c 160	28	2.7	279	10	AA605266	AA605266 no83c01.s
c 88	30	2.9	562	11	W05737	W05737 za88c04.r1	c 161	28	2.7	286	10	AA825946	AA825946 oe23f05.s
c 89	30	2.9	582	10	BE155132	BE155132 PMI-HT035	c 162	28	2.7	286	10	AA834766	AA834766 oe01d05.s
c 90	30	2.9	604	13	QA423585	QA423585 CITBI-E1	c 163	28	2.7	286	13	AA072686	AA072686 HS_2249.B
c 91	30	2.9	637	13	AQ936375	AQ936375 RPCI-11-S	c 164	28	2.7	288	11	F00153	F00153 HSB37E011.S
c 92	30	2.9	703	11	BG706072	BG706072 602669027	c 165	28	2.7	291	10	AA719769	AA719769 zh38f04.s
c 93	30	2.9	779	13	AQ739207	AQ739207 HS_5382.B	c 166	28	2.7	293	10	AW104788	AW104788 xd45c03.x
c 94	30	2.9	810	10	AL519355	AL519355 AL519355	c 167	28	2.7	295	11	BG978475	BG978475 CM2-CN006
c 95	30	2.9	867	10	AL519354	AL519354 AL519354	c 168	28	2.7	295	11	F31605	F31605 HSPD22913.H
c 96	30	2.9	868	10	BE544285	BE544285 601076760	c 169	28	2.7	306	13	AQ111420	AQ111420 CIT-HSP-2
c 97	30	2.9	1011	11	BF340406	BF340406 602036759	c 170	28	2.7	301	10	AQ095016	AQ095016 HS_3028.A
c 98	30	2.9	1020	11	BG615396	BG615396 602643974	c 171	28	2.7	311	13	AQ422892	AQ422892 CITBI-E1
c 99	29	2.8	234	13	AQ200492	AQ200492 RPCI11-63	c 172	28	2.7	312	10	AA228727	AA228727 nc46c05.r
c 100	29	2.8	249	11	BI005553	BI005553 CM0-HN020	c 173	28	2.7	313	10	AA846482	AA846482 ai85e08.s
c 101	29	2.8	261	11	BF877322	BF877322 QV2-ET010	c 174	28	2.7	318	10	AW517900	AW517900 xt66a04.x
c 102	29	2.8	318	10	AI432232	AI432232 th43c07.x	c 175	28	2.7	323	13	AQ487385	AQ487385 RPCI-11-2
c 103	29	2.8	352	10	AA515254	AA515254 ng69e10.s	c 176	28	2.7	324	11	H38416	H38416 yn82905.s1
c 104	29	2.8	377	11	H58407	H58407 yr20g12.r1	c 177	28	2.7	325	10	AA584183	AA584183 no11a07.s
c 105	29	2.8	391	10	AI865092	AI865092 wk07d06.x	c 178	28	2.7	326	10	AA886560	AA886560 ny42a09.s
c 106	29	2.8	392	13	AQ567558	AQ567558 HS_2118.B	c 179	28	2.7	327	10	AA366543	AA366543 EST77510
c 107	29	2.8	393	13	AO114347	AO114347 CIT-HSP-2	c 180	28	2.7	328	11	BG960534	BG960534 PM2-CT080
c 108	29	2.8	397	11	R33826	R33826 yH78g09.s1	c 181	28	2.7	330	13	AQ504662	AQ504662 RPCI-11-2
c 109	29	2.8	407	10	AA122219	AA122219 zn29c03.s	c 182	28	2.7	331	10	AI472766	AI472766 tal3f05.x
c 110	29	2.8	417	10	AW021007	AW021007 df17f06.y	c 183	28	2.7	333	11	BF763423	BF763423 CM1-CS003
c 111	29	2.8	437	10	BE065900	BE065900 KC2-BT031	c 184	28	2.7	333	13	AQ581730	AQ581730 RPCI-11-4
c 112	29	2.8	448	10	AA837335	AA837335 od68h02.s	c 185	28	2.7	339	13	AQ047180	AQ047180 RPCI11-35
c 113	29	2.8	459	10	BE065895	BE065895 KC2-BT031	c 186	28	2.7	348	10	AI344962	AI344962 tb01d09.x
c 114	29	2.8	463	10	BE065796	BE065796 KC2-BT031	c 187	28	2.7	351	10	AA652978	AA652978 AV652978
c 115	29	2.8	469	11	BF229305	BF229305 PM4-AN009	c 188	28	2.7	353	10	AA347203	AA347203 EST53463
c 116	29	2.8	469	11	BF229310	BF229310 PM4-AN009	c 189	28	2.7	354	11	H4786	H4786 yP20g12.r1
c 117	29	2.8	470	13	B80116	B80116 CIT-HSP-204	c 190	28	2.7	356	10	AA304577	AA304577 EST175803
c 118	29	2.8	475	13	B63124	B63124 CIT978SK-A-	c 191	28	2.7	358	10	BE175609	BE175609 RC5-HT058
c 119	29	2.8	477	10	BE065817	BE065817 RC2-BT031	c 192	28	2.7	358	13	AQ451866	AQ451866 HS_5183.A
c 120	29	2.8	491	11	T77730	T77730 yd16a11.s1	c 193	28	2.7	359	10	BE042026	BE042026 hj13f12.y
c 121	29	2.8	497	10	AW03196	AW03196 xdx3f09.x	c 194	28	2.7	361	10	BE043735	BE043735 hk42g06.y
c 122	29	2.8	500	13	AQ548911	AQ548911 RPCI-11-4	c 195	28	2.7	361	10	BE175589	BE175589 RC5-HT058
c 123	29	2.8	529	13	AO118113	AO118113 HS_3001.A	c 196	28	2.7	362	10	AA302964	AA302964 KC2-HT025
c 124	29	2.8	541	11	BF887415	BF887415 QV3-TN016	c 197	28	2.7	364	10	BE149242	BE149242 RC2-HT025
c 125	29	2.8	558	13	B68229	B68229 CIT978SK-A-	c 198	28	2.7	366	10	AA577815	AA577815 nn24d05.s
c 126	29	2.8	560	13	AQ937863	AQ937863 NB6-591C	c 199	28	2.7	366	13	AQ541818	AQ541818 RPCI-11-3
c 127	29	2.8	570	10	AI742503	AI742503 wq43f11.x	c 200	28	2.7	368	11	H00437	H00437 yJ24a02.s1
c 128	29	2.8	572	10	AW939304	AW939304 QV1-DT007	c 201	28	2.7	370	13	AQ695877	AQ695877 HS_5491.B
c 129	29	2.8	574	10	AL042377	AL042377 DXZ2P434A	c 202	28	2.7	377	11	BG988097	BG988097 MR2-HT116
c 130	29	2.8	590	13	AQ543668	AQ543668 RPCI-11-3	c 203	28	2.7	379	10	AI760985	AI760985 wh97a10.x
c 131	29	2.8	602	10	BE147684	BE147684 RPCI-11-3	c 204	28	2.7	381	13	AQ027138	AQ027138 CIT-HSP-2
c 132	29	2.8	649	13	AQ544103	AQ544103 RPCI-11-3	c 205	28	2.7	384	13	AQ199309	AQ199309 RPCI111-58
c 133	29	2.8	670	11	BF832615	BF832615 PM0-HT091	c 206	28	2.7	385	11	BG001009	BG001009 QV0-GN021
c 134	29	2.8	767	13	AQ742589	AQ742589 HS_5384.B	c 207	28	2.7	386	11	BG005228	BG005228 CM2-GN022
c 135	28	2.7	113	10	AA728909	AA728909 nw01a04.s	c 208	28	2.7	391	13	AQ234691	AQ234691 HS_2045.A
c 136	28	2.7	128	13	AQ471829	AQ471829 CITBI-E1-	c 209	28	2.7	394	13	B46875	B46875 HS-1066-A1-
c 137	28	2.7	129	10	AI478685	AI478685 tm01b07.x	c 210	28	2.7	395	13	AQ541822	AQ541822 RPCI-11-3
c 138	28	2.7	136	11	BI000436	BI000436 MR3-HN006	c 211	28	2.7	397	10	AI348892	AI348892 tb05d09.x
c 139	28	2.7	150	10	AA491440	AA491440 ab01a12.s	c 212	28	2.7	397	11	BG005129	BG005129 CM2-GN022
c 140	28	2.7	170	11	BG995538	BG995538 CM0-HT129	c 213	28	2.7	400	13	AQ0771299	AQ0771299 HS_2106.A
c 141	28	2.7	190	10	AA346883	AA346883 EST53099	c 214	28	2.7	401	10	AI130838	AI130838 qc09h11.x
c 142	28	2.7	205	13	B81851	B81851 RPCI11-17L4	c 215	28	2.7	402	10	AA548438	AA548438 nk03a11.s
c 143	28	2.7	208	10	AI266021	AI266021 qp63b10.x	c 216	28	2.7	410	11	R19998	R19998 yG38h06.r1
c 144	28	2.7	208	10	AK301555	AK301555 xs78f07.x	c 217	28	2.7	414	11	N27362	N27362 yW52g05.s1
c 145	28	2.7	209	11	BF997387	BF997387 QV3-GN020	c 218	28	2.7	415	13	AQ529967	AQ529967 RPCI-11-3
c 146	28	2.7	219	10	AA910123	AA910123 oe28h02.s	c 219	28	2.7	416	10	AA846966	AA846966 oe12d07.s
c 147	28	2.7	219	13	AZ693711	AZ693711 AST-2HLB4	c 220	28	2.7	416	10	BE155099	BE155099 PMI-HT035
c 148	28	2.7	223	10	AA288919	AA288919 od81a04.s	c 221	28	2.7	417	13	AQ481593	AQ481593 RPCI-11-2
c 149	28	2.7	231	10	AA298532	AA298532 EST114149	c 222	28	2.7	417	11	BG957324	BG957324 IL3-CT067
c 150	28	2.7	234	10	AA296569	AA296569 EST11094	c 223	28	2.7	420	11	R18066	R18066 yg15e11.r1
c 151	28	2.7	242	11	BF343920	BF343920 602015607	c 224	28	2.7	422	10	AA729755	AA729755 nx61d06.s
c 152	28	2.7	253	13	B53923	B53923 CIT-HSP-201	c 225	28	2.7	422	11	BF887129	BF887129 PM4-TN017
c 153	28	2.7	255	11	F19609	F19609 HSPD04214.H	c 226	28	2.7	422	11	BF929547	BF929547 RC1-MT016
c 154	28	2.7	265	13	AQ508343	AQ508343 RPCI-11-2	c 227	28	2.7	422	11	BF929547	BF929547 QV3-MT021
c 155	28	2.7	269	11	T20109	T20109 B466R Heart	c 228	28	2.7	426	10	AA551962	AA551962 nj83d08.s
c 156	28	2.7	273	13	AQ667078	AQ667078 HS_2107.A	c 229	28	2.7	426	13	AQ552478	AQ552478 RPCI-11-3
c 157	28	2.7	275	11	T69571	T69571 yc44b09.s1	c 230	28	2.7	431	11	BF836915	BF836915 CM3-HT099
c 158	28	2.7	275	11	BF847690	BF847690 IL5-EN008	c 231	28	2.7	431	10	AV763196	AV763196 AV763196

232	28	2.7	432	11	H46988	H46988	yo16c05.s1	c 305	28	2.7	582	13	AQ264917	CITBI-El-
233	28	2.7	433	10	AW277123	AW277123	xp63b07.x	306	28	2.7	584	11	BI010689	MR2-EN009
234	28	2.7	434	13	AQ043528	AQ043528	CIT-HSP-2	307	28	2.7	586	13	AQ190336	RPCI-11-1
235	28	2.7	439	10	AV732618	AV732618	AV732618	c 308	28	2.7	597	13	AQ571789	HS-2094.A
236	28	2.7	439	11	BF798467	BF798467	CMO-CI005	c 309	28	2.7	612	11	BF980214	602288048
237	28	2.7	440	11	N22058	N22058	yy45h03.s1	c 310	28	2.7	618	13	B95496	CIT-HSP-216
238	28	2.7	441	10	AA325466	AA325466	ES278454	c 311	28	2.7	618	13	AQ543062	RPCI-11-3
239	28	2.7	441	10	AA325466	AA325466	ES278454	c 312	28	2.7	622	13	AQ017432	CIT-HSP-2
240	28	2.7	444	10	BE000907	BE000907	RC2-BN007	c 313	28	2.7	622	13	AQ017432	CITBI-El-
241	28	2.7	448	11	H19684	H19684	ym59f10.s1	c 314	28	2.7	644	10	AI207582	HA2946.Hu
242	28	2.7	450	10	AI423231	AI423231	tf26f12.x	c 315	28	2.7	650	13	AQ425110	RPCI-11-2
243	28	2.7	450	13	AQ267366	AQ267366	RPCI11-68	c 316	28	2.7	679	13	AQ321467	RPCI11-92
244	28	2.7	451	10	AI732074	AI732074	nk03a11.x	c 317	28	2.7	702	13	AQ478621	RPCI-11-2
245	28	2.7	451	11	BE887225	BE887225	601508536	c 318	28	2.7	724	10	AL048376	AL048376
246	28	2.7	458	11	BF887098	BF887098	PM4-TN017	c 319	28	2.7	736	11	BG482301	602526637
247	28	2.7	459	10	AI820988	AI820988	nk03a11.y	c 320	28	2.7	755	11	BG481747	BG481747
248	28	2.7	460	10	AA603502	AA603502	np07c01.s	c 321	28	2.7	759	13	AQ538545	RPCI-11-3
249	28	2.7	460	11	R92256	R92256	YQ06e01.r1	c 322	28	2.7	797	11	BG187463	RST6456.A
250	28	2.7	460	13	AA015971	AA015971	ze30g09.f	c 323	28	2.7	804	11	BG428355	602495340
251	28	2.7	462	10	BE150018	BE150018	QV3-HT026	c 324	28	2.7	805	10	AL120197	AU120197
252	28	2.7	462	10	BE150018	BE150018	QV3-HT026	c 325	28	2.7	828	11	BG421216	602496201
253	28	2.7	462	11	BF914803	BF914803	IL3-UT011	c 326	28	2.7	855	10	AL157580	AL157580
254	28	2.7	464	10	AI338899	AI338899	qql9e07.x	c 327	28	2.7	861	10	AI525789	PTI.3-05-
255	28	2.7	465	10	BE138593	BE138593	x779l12.x	c 328	28	2.7	879	13	AQ741630	AQ741630
256	28	2.7	465	13	AQ178738	AQ178738	HS-2256.A	c 329	28	2.7	885	11	BF575954	BF575954
257	28	2.7	466	11	BF926463	BF926463	QV3-NT021	c 330	28	2.7	895	11	BI091119	602854730
258	28	2.7	467	13	AQ167841	AQ167841	ni61f10.s	c 331	28	2.7	911	11	BE910447	BE910447
259	28	2.7	467	10	AA525927	AA525927	ni61f10.s	c 332	28	2.7	934	13	AQ750230	BE910447
260	28	2.7	475	11	BF770393	BF770393	RC2-IT004	c 333	28	2.7	946	10	BE621403	BE621403
261	28	2.7	475	13	AQ783476	AQ783476	HS-3121.A	c 334	28	2.7	959	11	BG340032	BG340032
262	28	2.7	484	11	BG285245	BG285245	602409568	c 335	28	2.7	1048	11	BF340593	BF340593
263	28	2.7	485	13	AQ550476	AQ550476	RPCI-11-4	c 336	28	2.7	1129	11	BF340593	BF340593
264	28	2.7	486	11	BI003340	BI003340	PMO-HN007	c 337	28	2.7	1606	13	AF101950	AF101950
265	28	2.7	488	11	BF843512	BF843512	RC1-HT022	c 338	28	2.7	2225	12	AF130097	Homo.sapi
266	28	2.7	490	11	BF843512	BF843512	RC1-HT022	c 339	27	2.6	105	11	BG951123	CMO-CIT069
267	28	2.7	491	10	AL044807	AL044807	DKEZp340	c 340	27	2.6	105	11	BI041679	IL3-NT029
268	28	2.7	493	13	B36003	B36003	HS-1031-A2-	c 341	27	2.6	121	11	BF897203	IL3-NT029
269	28	2.7	495	10	AL044445	AL044445	DKEZp340	c 342	27	2.6	134	11	BF880108	IL3-NT029
270	28	2.7	498	10	AI472093	AI472093	tj80e07.x	c 343	27	2.6	142	10	BE180038	IL3-NT029
271	28	2.7	500	13	AQ471002	AQ471002	CITBI-El-	c 344	27	2.6	149	11	BG999334	IL3-NT029
272	28	2.7	501	10	AA074774	AA074774	zm84a10.s	c 345	27	2.6	149	13	AQ533020	PMO-HN007
273	28	2.7	501	13	AQ261648	AQ261648	CITBI-El-	c 346	27	2.6	153	11	AG857779	PMO-HN007
274	28	2.7	514	13	AQ175560	AQ175560	HS-3215.A	c 347	27	2.6	160	10	AG857779	PMO-HN007
275	28	2.7	517	10	AI379712	AI379712	tc41a06.x	c 348	27	2.6	161	11	BG949998	PMO-HN007
276	28	2.7	521	10	AW936276	AW936276	QV0-DT002	c 349	27	2.6	162	11	BF878268	CM3-ET009
277	28	2.7	521	13	AQ122627	AQ122627	HS-3086.A	c 350	27	2.6	165	10	BF878268	CM3-ET009
278	28	2.7	523	11	BI029851	BI029851	MR4-MT025	c 351	27	2.6	167	10	AW857759	RC2-CIT032
279	28	2.7	523	13	B69012	B69012	CIT-HSP-202	c 352	27	2.6	177	10	AA658018	nul15c02.s
280	28	2.7	527	13	AQ184178	AQ184178	HS-3218.A	c 353	27	2.6	180	11	BI038819	IL2-UM007
281	28	2.7	528	13	AQ321460	AQ321460	RPCI11-92	c 354	27	2.6	183	11	BF0909161	CM3-NT028
282	28	2.7	531	10	AU158549	AU158549	HS-3121.A	c 355	27	2.6	190	10	AI349649	CM3-NT028
283	28	2.7	531	13	AQ783497	AQ783497	MR1-RT004	c 356	27	2.6	199	13	AQ423011	CITBI-El-
284	28	2.7	539	13	AQ783497	AQ783497	MR1-RT004	c 357	27	2.6	200	10	AA157711	zo46e03.s
285	28	2.7	542	11	BF821018	BF821018	RC2-BN013	c 358	27	2.6	200	10	BE158821	IL2-HT039
286	28	2.7	543	13	AQ483630	AQ483630	RC2-CI008	c 359	27	2.6	200	10	BE158821	IL2-HT039
287	28	2.7	543	13	B15366	B15366	344013.TV.C	c 360	27	2.6	201	11	BI029152	IL5-MT026
288	28	2.7	544	10	AW512773	AW512773	xt77e04.x	c 361	27	2.6	201	13	B42128	IL5-MT026
289	28	2.7	544	11	BF737356	BF737356	PM4-KT004	c 362	27	2.6	208	11	BF853074	HS-1055-A2-
290	28	2.7	544	13	AQ213301	AQ213301	HS-3250.B	c 363	27	2.6	208	11	BF853074	MR2-EN009
291	28	2.7	545	13	AQ491848	AQ491848	RPCI-11-2	c 364	27	2.6	211	10	AA605074	no74d02.s
292	28	2.7	546	13	AQ410315	AQ410315	HS-5111.A	c 365	27	2.6	211	13	BI01997	CSU-145011
293	28	2.7	546	13	AQ493701	AQ493701	HS-5064.B	c 366	27	2.6	215	13	B55313	CIT-HSP-386
294	28	2.7	547	10	AA147076	AA147076	zo32g02.s	c 367	27	2.6	219	10	AA2293322	no74d02.s
295	28	2.7	549	10	BE006648	BE006648	RC2-BN013	c 368	27	2.6	220	10	BE158395	CSU-145011
296	28	2.7	550	13	AQ0701930	AQ0701930	HS-5396.B	c 369	27	2.6	222	13	BE158395	CIT-HSP-386
297	28	2.7	556	11	BF812313	BF812313	RC2-CI008	c 370	27	2.6	223	10	AQ262194	CMO-HT038
298	28	2.7	557	11	N63823	N63823	za26d10.s1	c 371	27	2.6	227	13	BE158833	CMO-HT038
299	28	2.7	562	13	AQ496544	AQ496544	HS-5134.B	c 372	27	2.6	227	13	BE158833	CMO-HT038
300	28	2.7	563	13	AQ456563	AQ456563	HS-5154.B	c 373	27	2.6	228	13	AQ553065	CMO-HT038
301	28	2.7	570	13	AQ311678	AQ311678	RPCI11-11	c 374	27	2.6	230	13	AQ40316	CMO-HT038
302	28	2.7	579	11	AA206366	AA206366	zo58b07.s	c 375	27	2.6	233	10	BE261473	CMO-HT038
303	28	2.7	579	11	BF340510	BF340510	602037142	c 376	27	2.6	235	11	TS1084	BB261473
304	28	2.7	580	11	BF925837	BF925837	RC3-NT015	c 377	27	2.6	239	11	TS1084	BB261473
	28	2.7	582	10	AA577696	AA577696	nn22h03.s	377	27	2.6	240	13	B37604	BB261473

C 378 27 2.6 241 11 BI014567  
C 379 27 2.6 244 11 BI038137  
C 380 27 2.6 244 13 AQ275929  
C 381 27 2.6 245 11 BI038809  
C 382 27 2.6 247 10 AW170716  
C 383 27 2.6 248 10 AA730916  
C 384 27 2.6 249 11 BF772982  
C 385 27 2.6 250 10 AA482556  
C 386 27 2.6 253 10 BA649322  
C 387 27 2.6 254 10 AI567372  
C 388 27 2.6 255 10 AW183015  
C 389 27 2.6 256 10 AI367756  
C 390 27 2.6 257 10 AI267161  
C 391 27 2.6 257 13 B87164  
C 392 27 2.6 258 13 AQ626001  
C 393 27 2.6 258 10 AA782405  
C 394 27 2.6 259 10 AI540275  
C 395 27 2.6 261 10 AI371864  
C 396 27 2.6 262 13 AQ106193  
C 397 27 2.6 263 10 AA438575  
C 398 27 2.6 264 10 AA229255  
C 399 27 2.6 265 10 AL134042  
C 400 27 2.6 266 10 AA658320  
C 401 27 2.6 266 11 BG980642  
C 402 27 2.6 267 10 AW270693  
C 403 27 2.6 268 10 AI758866  
C 404 27 2.6 269 10 AI332831  
C 405 27 2.6 269 10 AA113129  
C 406 27 2.6 269 10 AI719747  
C 407 27 2.6 269 11 R02101  
C 408 27 2.6 270 11 T26442  
C 409 27 2.6 271 10 AI269212  
C 410 27 2.6 271 10 AI761458  
C 411 27 2.6 271 10 BE158836  
C 412 27 2.6 272 11 N64587  
C 413 27 2.6 273 10 AI473842  
C 414 27 2.6 273 10 AV739829  
C 415 27 2.6 277 11 H95865  
C 416 27 2.6 278 11 BI010361  
C 417 27 2.6 278 11 M78108  
C 418 27 2.6 281 10 AA984754  
C 419 27 2.6 283 13 AQ593014  
C 420 27 2.6 285 13 AQ017472  
C 421 27 2.6 285 13 AQ017472  
C 422 27 2.6 286 10 AL042630  
C 423 27 2.6 287 10 AA984829  
C 424 27 2.6 288 11 T39156  
C 425 27 2.6 291 10 AA87892  
C 426 27 2.6 292 13 AQ086725  
C 427 27 2.6 293 13 B85662  
C 428 27 2.6 294 11 BI036077  
C 429 27 2.6 296 10 AA856945  
C 430 27 2.6 297 10 AA883910  
C 431 27 2.6 297 10 AA244303  
C 432 27 2.6 297 11 N70847  
C 433 27 2.6 297 13 AQ056514  
C 434 27 2.6 298 11 BI016488  
C 435 27 2.6 298 13 A2518668  
C 436 27 2.6 299 10 AA297207  
C 437 27 2.6 301 11 F32823  
C 438 27 2.6 303 10 AI697466  
C 439 27 2.6 304 10 AW731919  
C 440 27 2.6 305 10 AA662452  
C 441 27 2.6 308 11 T52822  
C 442 27 2.6 309 11 H24040  
C 443 27 2.6 311 10 AA158983  
C 444 27 2.6 312 10 AA584810  
C 445 27 2.6 312 10 AA342457  
C 446 27 2.6 313 10 AA516028  
C 447 27 2.6 313 10 AW834482  
C 448 27 2.6 313 10 AA565555  
C 449 27 2.6 314 10 AA218850  
C 450 27 2.6 314 11 BI044844

BI014567 CM3-ET009  
BI038137 CM3-NT028  
AQ275929 CITBI-E1-  
BI038809 CM3-NT028  
AW170716 xi97a08.x  
AA730916 nw50e01.s  
BF772982 CM3-IT004  
AA482556 zt34c06.s  
BE464922 hs88b06.x  
AI567372 tp67b09.x  
AW183015 xp99g07.x  
AI367756 qt55g08.x  
AI267161 aq49f02.x  
B87164 RPCI11-27P2  
AQ626001 CITBI-E1-  
AA782405 ai31h08.s  
AI540275 t46h05.x  
AI371864 ta34h06.x  
AQ106193 HS\_3055\_A  
AA58575 aa83h02.s  
AA229255 nc47e04.s  
AL134042 DKFzp547B  
AA658320 nul7e05.s  
BG980642 CM4-CN008  
AW270693 xs56r12.x  
AI758866 t94a06.x  
AI332831 qp96f09.x  
AA113129 zm29d07.s  
AI719747 as92d09.x  
R02101 ye83c07.s1  
T26442 AB211H9F in  
AI269212 qj68g04.x  
AI761458 wq65h07.x  
BE158836 IJ2-HT039  
N64587 yz51h09.s1  
AI473842 tcl7f01.x  
AV739829 AV739829  
H95865 yw48c10.s1  
BI010361 MR2-EN009  
M78108 EST01698 Su  
AA984754 am9f04.s  
AQ593014 HS\_5430\_B  
AQ017472 CIT-HSP-2  
AO549112 RPCI-11-4  
AL042630 DKFz434M  
AA984829 am93f09.s  
T39156 ya01c10.r2  
AA487892 abl2b07.s  
AQ086725 HS\_2188\_A  
B85662 RPCI11-20H1  
BI036077 IJ2-NT020  
AA856945 oe03c08.s  
AA883910 aj13a10.s  
AA244303 nc14b01.r  
N70847 za88a12.s1  
AQ056514 HS\_5386\_A  
BI016488 PM0-ET025  
A2518668 RPCI-11-2  
AA297207 EST112750  
F32823 HSPD25866 H  
AI697466 tq08q07.x  
AW731919 ba03a11.x  
AA662452 nu91f04.s  
T52822 ya81a11.s1  
H24040 ym34a07.s1  
AA158983 zo57a05.s  
AA584810 nol0d08.s  
AA342457 EST48183  
AA516028 ng65a11.s  
AW834482 MR2-TT001  
AA565555 nk42e06.s  
AA218850 zq15b09.s  
BI044844 PM0-OT020

C 451 27 2.6 314 11 N85485  
C 452 27 2.6 316 10 AW772696  
C 453 27 2.6 318 10 AA663420  
C 454 27 2.6 318 10 AA446804  
C 455 27 2.6 318 11 BG006858  
C 456 27 2.6 320 13 AQ392145  
C 457 27 2.6 321 13 AQ378288  
C 458 27 2.6 321 13 AQ388603  
C 459 27 2.6 324 13 AQ14621  
C 460 27 2.6 325 13 B91498  
C 461 27 2.6 327 10 AA985332  
C 462 27 2.6 328 10 AA632640  
C 463 27 2.6 329 10 AI284467  
C 464 27 2.6 331 11 BF886839  
C 465 27 2.6 333 10 AA300070  
C 466 27 2.6 335 11 F012742  
C 467 27 2.6 338 11 F01304  
C 468 27 2.6 338 11 BF742465  
C 469 27 2.6 339 13 AQ081573  
C 470 27 2.6 340 10 AA748140  
C 471 27 2.6 340 10 AA604632  
C 472 27 2.6 341 11 BF899031  
C 473 27 2.6 342 10 AA196204  
C 474 27 2.6 343 13 AQ043469  
C 475 27 2.6 344 11 BG958931  
C 476 27 2.6 345 11 BF873595  
C 477 27 2.6 349 13 AQ603977  
C 478 27 2.6 350 10 AI475147  
C 479 27 2.6 350 11 F08649  
C 480 27 2.6 354 10 AW179336  
C 481 27 2.6 355 10 AA747957  
C 482 27 2.6 356 10 AI246884  
C 483 27 2.6 357 10 BE069060  
C 484 27 2.6 357 13 AQ239144  
C 485 27 2.6 359 10 BE068949  
C 486 27 2.6 363 10 AW812826  
C 487 27 2.6 364 13 AQ047409  
C 488 27 2.6 367 11 BI031199  
C 489 27 2.6 367 11 W21392  
C 490 27 2.6 369 10 AA987606  
C 491 27 2.6 370 10 AA653041  
C 492 27 2.6 371 10 AI734052  
C 493 27 2.6 372 10 AW069769  
C 494 27 2.6 373 10 AA229034  
C 495 27 2.6 373 10 AA573631  
C 496 27 2.6 373 11 BF922719  
C 497 27 2.6 374 10 AJ346511  
C 498 27 2.6 374 10 AA533855  
C 499 27 2.6 375 10 AV710483  
C 500 27 2.6 375 11 BG942919  
C 501 27 2.6 375 11 H92152  
C 502 27 2.6 375 11 BF919050  
C 503 27 2.6 376 13 AQ281592  
C 504 27 2.6 378 10 BA837716  
C 505 27 2.6 378 10 BE562811  
C 506 27 2.6 379 11 BG170207  
C 507 27 2.6 381 10 AW793601  
C 508 27 2.6 382 10 AA678632  
C 509 27 2.6 382 10 AI201706  
C 510 27 2.6 382 13 AQ588377  
C 511 27 2.6 383 10 AA329122  
C 512 27 2.6 384 10 AI354409  
C 513 27 2.6 384 11 BG944366  
C 514 27 2.6 384 11 BG944366  
C 515 27 2.6 384 11 BG980850  
C 516 27 2.6 384 13 AQ208401  
C 517 27 2.6 386 13 AQ263512  
C 518 27 2.6 387 10 AA502098  
C 519 27 2.6 387 10 AA506838  
C 520 27 2.6 389 10 AW166641  
C 521 27 2.6 390 11 T69896  
C 522 27 2.6 391 10 AA400436  
C 523 27 2.6 392 11 BF849841

N85485 J3701F Huma  
AW772696 hn05g04.x  
AA663420 ac51c07.s  
AA446804 zw80g04.r  
BG006858 IL5-GN023  
AQ392145 CITBI-E1-  
AQ378288 RPCI11-15  
AQ388603 RPCI11-15  
AQ14621 RPCI11-10  
B91498 CIT-HSP-217  
AA985332 am79g05.s  
AA632640 np87b03.s  
AI284467 qh82a03.x  
BF886839 PM3-TN018  
AA300070 oc4a10.s  
BG012742 IL5-GN023  
F01304 HSBG6E012 S  
BF742465 RC1-BT025  
AQ081573 RPCI11-34  
AA748140 nx79h10.s  
AA604632 n084c10.s  
BF899031 PM2-MT020  
AA196204 zp92c04.s  
AQ043469 CIT-HSP-2  
BG958931 PM2-CT026  
BF873595 IL5-ET011  
AQ603977 HS\_2136\_A  
AI475147 t179a09.x  
F08649 HSCIDE111 n  
AW179336 PM4-ST012  
AA747957 nx68d12.s  
AI246884 qx51e12.x  
BE069060 QV3-BT037  
AQ239144 RPCI11-75  
BE068949 RC5-BT037  
AW812826 RC3-ST018  
AQ047409 CIT-HSP-2  
BI031199 IL5-MT020  
W21392 zb9a09.r1  
AA987606 of93a11.s  
AA653041 ns71e11.s  
AI734052 nc09f07.y  
AW069769 ct50c02.x  
AA229034 nc50a10.s  
AA573631 nf43h05.s  
BF922719 OV4-NT025  
AJ346511 SHr-00001  
AA533855 nj94f04.s  
AV710483 AV710483  
BG942919 ax31c04.x  
H92152 ys84e09.r1  
BF919050 CM0-NT013  
AQ281592 RPCI11-88  
AA837716 oe06d05.s  
BE562811 601336517  
BG170207 602321718  
AW793601 MK1-UM000  
AA678632 an07a02.s  
AI201706 q144c10.x  
AQ588377 CITBI-E1-  
AA329122 EST32968  
AI354409 qul7a02.x  
AA515926 ng64g01.s  
BG944366 ax49f07.x  
BG980850 CM4-CN008  
AQ208401 HS\_3230\_A  
AQ263512 CITBI-E1-  
AA502098 ng58a03.s  
AA506838 nh62f07.s  
AW166641 xg54g09.x  
T69896 YC18h04.r1  
AA400436 zu62b12.s  
BF849841 PM4-EN006



524	27	2.6	393	10	AW265614	AW265614	xq72e08.x	c 597	27	2.6	436	13	AQ276331	AQ276331	CITBI-El-
525	27	2.6	393	11	BG982190	BG982190	CM4-CN008	c 598	27	2.6	436	13	AQ511172	AQ511172	HS_5181_B
526	27	2.6	393	13	AQ191341	AQ191341	HS_3234_A	c 599	27	2.6	437	10	AA991792	AA991792	HS_5181_B
527	27	2.6	394	13	AQ074696	AQ074696	CIT-HSP-2	600	27	2.6	437	10	A1636336	A1636336	CIT-HSP-2
528	27	2.6	395	11	W45412	W45412	zc82b04.s1	601	27	2.6	438	10	AA188421	AA188421	zc82b04.s1
529	27	2.6	397	11	AA659421	AA659421	nu23e12.s	c 602	27	2.6	438	11	AA197174	AA197174	zc82b04.s1
530	27	2.6	397	11	H79436	H79436	yu28c04.r1	603	27	2.6	438	11	T93089	T93089	zc82b04.s1
531	27	2.6	398	13	AQ261641	AQ261641	CITBI-El-	604	27	2.6	438	13	AQ040754	AQ040754	CIT-HSP-2
532	27	2.6	399	11	H08159	H08159	Y187d04.s1	c 605	27	2.6	439	10	AA011026	AA011026	CIT-HSP-2
533	27	2.6	400	10	AA311535	AA311535	EST182278	c 606	27	2.6	439	13	AQ936180	AQ936180	RPCI-11-S
534	27	2.6	401	13	AQ012471	AQ012471	CIT-HSP-2	c 607	27	2.6	440	10	AA300177	AA300177	EST18206
535	27	2.6	402	11	BF933868	BF933868	RC5-NTP018	c 608	27	2.6	440	10	AA306177	AA306177	EST18206
536	27	2.6	403	10	AA631729	AA631729	np79e03.s	c 609	27	2.6	440	13	AQ462114	AQ462114	HS_5118_B
537	27	2.6	403	10	AA226144	AA226144	nc09f07.r	610	27	2.6	441	10	AA410715	AA410715	f008a02.y
538	27	2.6	403	11	BI063253	BI063253	IL3-UT011	c 611	27	2.6	441	11	N72479	N72479	YV42d12.r1
539	27	2.6	405	10	AA199748	AA199748	zq76e08.s	c 612	27	2.6	442	11	N72225	N72225	zq76e08.s
540	27	2.6	405	11	T84567	T84567	yq50a06.r1	c 613	27	2.6	442	13	AQ819222	AQ819222	HS_5300_A
541	27	2.6	405	11	BF901439	BF901439	IL2-MT017	c 614	27	2.6	443	10	AL603359	AL603359	DKFZp686C
542	27	2.6	406	10	AW959130	AW959130	EST371200	c 615	27	2.6	443	13	AQ0702809	AQ0702809	HS_5449_B
543	27	2.6	406	11	W73597	W73597	zd55c04.s1	c 616	27	2.6	444	13	AQ013442	AQ013442	RPCI-11-25
544	27	2.6	407	10	AI583532	AI583532	ts12c10.x	c 617	27	2.6	445	10	AT800457	AT800457	t114f08.x
545	27	2.6	407	10	AV755184	AV755184	AV755184	618	27	2.6	447	10	AA587385	AA587385	nm82908.s
546	27	2.6	409	11	H65832	H65832	Yr69d02.s1	619	27	2.6	447	10	AQ062147	AQ062147	CIT-HSP-2
547	27	2.6	409	11	H79260	H79260	YU48b06.s1	620	27	2.6	447	13	AQ193626	AQ193626	CIT-HSP-2
548	27	2.6	410	13	B55380	B55380	CIT-HSP-386	c 621	27	2.6	447	13	AQ393777	AQ393777	CITBI-El-
549	27	2.6	410	13	AQ277094	AQ277094	CITBI-El-	c 622	27	2.6	448	11	H11242	H11242	Ym09b10.s1
550	27	2.6	411	10	AA458745	AA458745	aa87d05.r	c 623	27	2.6	448	11	BF727198	BF727198	BY18b06.y
551	27	2.6	412	10	AI167842	AI167842	oz90g07.x	c 624	27	2.6	448	11	BF727199	BF727199	BY18b06.y
552	27	2.6	413	10	AA729070	AA729070	nm3b01.s	625	27	2.6	448	13	AQ425680	AQ425680	CITBI-El-
553	27	2.6	413	10	AL597195	AL597195	DKFZp313M	c 626	27	2.6	449	13	AQ562747	AQ562747	HS_2079_B
554	27	2.6	413	11	H78610	H78610	YU26b04.s1	c 627	27	2.6	451	13	AQ318988	AQ318988	RPCI-11-1
555	27	2.6	413	13	AQ167176	AQ167176	HS_3150_A	c 628	27	2.6	451	13	BE207952	BE207952	BS50a12.x
556	27	2.6	414	10	AA100281	AA100281	z084f02.s	c 629	27	2.6	453	13	AQ178863	AQ178863	HS_3061_B
557	27	2.6	414	10	AA173391	AA173391	zo64c12.s	c 630	27	2.6	454	13	AQ178863	AQ178863	HS_3061_B
558	27	2.6	415	10	AA555165	AA555165	n116e05.s	c 631	27	2.6	455	10	A1089524	A1089524	oz22a01.x
559	27	2.6	415	10	AA714282	AA714282	nv35c08.s	c 632	27	2.6	457	13	AQ084913	AQ084913	HS_2267_B
560	27	2.6	415	10	AA083046	AA083046	nv35c08.s	c 633	27	2.6	458	10	AT471827	AT471827	MR4-NT014
561	27	2.6	415	10	AA187277	AA187277	zp59a01.s	c 634	27	2.6	459	11	BF923177	BF923177	MR4-NT014
562	27	2.6	415	10	AA226142	AA226142	nc09f05.r	c 635	27	2.6	461	11	BF855142	BF855142	MR4-NT011
563	27	2.6	415	10	AA229080	AA229080	nc50d10.r	c 636	27	2.6	463	10	AA668863	AA668863	aa80e09.s
564	27	2.6	417	10	AI285453	AI285453	qt64a12.x	c 637	27	2.6	464	10	AA128941	AA128941	zo11c09.s
565	27	2.6	417	10	AA487475	AA487475	ab23c04.s	c 638	27	2.6	464	10	BE301503	BE301503	bb74d06.x
566	27	2.6	417	10	AA501384	AA501384	ne66e02.s	c 639	27	2.6	464	10	AA584067	AA584067	no04f03.s
567	27	2.6	418	10	AA075975	AA075975	zm75d02.r	c 640	27	2.6	466	10	AA670146	AA670146	ab65b01.s
568	27	2.6	418	10	AA225949	AA225949	nc11f02.y	c 641	27	2.6	468	11	BI019022	BI019022	IL3-MT026
569	27	2.6	418	10	AA452998	AA452998	zx41c08.r	c 642	27	2.6	468	13	BF7063	BF7063	CIT-HSP-201
570	27	2.6	418	13	AQ480737	AQ480737	RPCI-11-2	c 643	27	2.6	471	11	AT191451	AT191451	qz32e08.s
571	27	2.6	419	10	BE206447	BE206447	bb64a12.x	c 644	27	2.6	471	13	BF808415	BF808415	RC3-C1004
572	27	2.6	419	13	AZ515766	AZ515766	RPCI-11-2	c 645	27	2.6	473	10	AQ472036	AQ472036	CITBI-El-
573	27	2.6	420	13	AQ210457	AQ210457	HS_2202_B	c 646	27	2.6	473	10	AT677655	AT677655	wc79g01.x
574	27	2.6	421	10	AT446708	AT446708	t198b08.x	c 647	27	2.6	473	13	AM873401	AM873401	h199a03.x
575	27	2.6	423	13	AQ615127	AQ615127	HS_5120_B	c 648	27	2.6	475	10	AA431791	AA431791	h199a03.x
576	27	2.6	425	11	W57877	W57877	zd17b09.s1	c 649	27	2.6	475	10	AA31791	AA31791	CIT978SK-17
577	27	2.6	425	13	AQ696891	AQ696891	HS_3526_A	c 650	27	2.6	475	11	BI063789	BI063789	zw80g04.s
578	27	2.6	426	13	AQ463860	AQ463860	HS_5080_A	c 651	27	2.6	477	10	AA019548	AA019548	aa34a05.r
579	27	2.6	426	10	AA181070	AA181070	zp67e03.s	c 652	27	2.6	477	10	AA481174	AA481174	aa34a05.r
580	27	2.6	426	11	BI030534	BI030534	IL5-UT011	c 653	27	2.6	478	11	BG009651	BG009651	MR4-GN019
581	27	2.6	426	11	BI062364	BI062364	IL3-UT011	c 654	27	2.6	479	10	A1002389	A1002389	or78a04.s
582	27	2.6	426	11	BI062364	BI062364	IL3-UT011	c 655	27	2.6	479	10	A1002389	A1002389	or78a04.s
583	27	2.6	428	10	AT049676	AT049676	aa33b04.x	c 656	27	2.6	479	11	T71765	T71765	zc23a06.x
584	27	2.6	429	11	BI028061	BI028061	IL5-MT020	c 657	27	2.6	479	13	AQ217365	AQ217365	Yc64a03.s1
585	27	2.6	429	11	BF902215	BF902215	IL5-MT020	c 658	27	2.6	480	10	AA584493	AA584493	HS_2190_A
586	27	2.6	431	10	AA954188	AA954188	op08b06.s	c 659	27	2.6	481	10	AA854632	AA854632	no11g07.s
587	27	2.6	431	10	AA305148	AA305148	xv89g01.x	c 660	27	2.6	481	11	BF149121	BF149121	a773g08.s
588	27	2.6	431	11	BI028066	BI028066	IL5-MT020	c 661	27	2.6	482	11	BF754469	BF754469	PM3-CT054
589	27	2.6	432	10	AW772774	AW772774	hl96d02.x	c 662	27	2.6	484	10	BE147634	BE147634	RC1-HT022
590	27	2.6	432	10	AA508485	AA508485	nh66e08.s	c 663	27	2.6	484	11	BF585939	BF585939	Yb62g01.s1
591	27	2.6	432	13	AQ015260	AQ015260	CIT-HSP-2	c 664	27	2.6	484	11	BF958535	BF958535	601645163
592	27	2.6	432	13	AQ015260	AQ015260	CIT-HSP-2	c 665	27	2.6	487	13	AQ336600	AQ336600	HS_5000_A
593	27	2.6	433	13	AQ146914	AQ146914	HS_2246_A	c 666	27	2.6	487	11	BG222269	BG222269	naf74b03.
594	27	2.6	433	10	AA143440	AA143440	aa55c10.s	c 667	27	2.6	488	10	AT800437	AT800437	HS_5250_B
595	27	2.6	433	10	AA226095	AA226095	nc11f02.r	c 668	27	2.6	488	13	A1800437	A1800437	t114d08.x
596	27	2.6	435	13	RQ202410	RQ202410	RPC111-58	c 669	27	2.6	489	13	AQ403521	AQ403521	HS_5055_B

670 27 AL044793 DKF2p434N  
671 27 AW471092 xu06h04.x  
672 27 AI500640 nt93e02.x  
c 673 27 AQ422506 RPCI-11-1  
674 27 BG777252 602664483  
675 27 BF919931 QVO-NT015  
c 676 27 B81834  
677 27 B81834  
678 27 AQ496808  
679 27 AQ620271  
680 27 AQ620271 HS\_5182\_B  
681 27 AQ226133 HS\_2015\_B  
c 682 27 AQ489571 RPCI-11-2  
683 27 AL597533 DKF2p313G  
c 684 27 AV739938 AV739938  
685 27 AA626478 ab53h12.f  
c 686 27 AQ429861 HS\_5077\_A  
687 27 AV739810 AV739810  
c 688 27 AA492114 ng17d03.s  
689 27 AA400544 zu62b12.f  
690 27 AQ211741 HS\_3239\_A  
c 691 27 AA670456 ae62h06.s  
692 27 AA482950 nf49e09.s  
c 693 27 AQ244398  
694 27 AQ526160 HS\_5312\_B  
c 695 27 AQ512553 HS\_5119\_B  
696 27 AQ562887  
c 697 27 AI732085  
698 27 AQ512745  
c 699 27 BE178481  
700 27 BE065970 RC3-HT060  
c 701 27 AW167073  
702 27 BF737749  
c 703 27 BI039053  
c 704 27 BI039054  
705 27 BG009072  
c 706 27 AU144449  
707 27 AV720925  
c 708 27 AQ536676 RPCI-11-3  
709 27 AQ691213 HS\_5338\_A  
c 710 27 BG056397 na6f68b04  
c 711 27 AQ078350 HS\_5569\_A  
c 712 27 AQ08037  
713 27 AL042731  
714 27 AL042731  
715 27 BI021436  
c 716 27 AQ480181  
c 717 27 AQ332193  
c 718 27 BI039969  
c 719 27 BF448499  
c 720 27 AW975367  
c 721 27 H23506  
c 722 27 AZ254674  
c 723 27 BF826816  
c 724 27 AW939214  
c 725 27 AQ570204  
c 726 27 AQ721483  
c 727 27 AQ352086  
c 728 27 AQ000610  
c 729 27 AQ020745  
c 730 27 BE740827  
c 731 27 BG983313  
c 732 27 N78063  
c 733 27 AQ678716  
c 734 27 BE704103  
c 735 27 AQ334955  
c 736 27 AQ533504  
c 737 27 BF675599  
c 738 27 BE143857  
c 739 27 BI0354690  
c 740 27 BF954499  
c 741 27 BE299730  
c 742 27 AQ441794 HS\_5106\_B  
c 743 27 AW603818 CN0-CN004  
c 744 27 AL044793 DKF2p434N  
c 745 27 AW471092 xu06h04.x  
c 746 27 AI500640 nt93e02.x  
c 747 27 AQ422506 RPCI-11-1  
c 748 27 BG777252 602664483  
c 749 27 BF919931 QVO-NT015  
c 750 27 B81834  
c 751 27 B81834  
c 752 27 AQ496808  
c 753 27 AQ620271  
c 754 27 AQ620271 HS\_5182\_B  
c 755 27 AQ226133 HS\_2015\_B  
c 756 27 AQ489571 RPCI-11-2  
c 757 27 AL597533 DKF2p313G  
c 758 27 AV739938 AV739938  
c 759 27 AA626478 ab53h12.f  
c 760 27 AQ429861 HS\_5077\_A  
c 761 27 AV739810 AV739810  
c 762 27 AA492114 ng17d03.s  
c 763 27 AA400544 zu62b12.f  
c 764 27 AQ211741 HS\_3239\_A  
c 765 27 AA670456 ae62h06.s  
c 766 27 AA482950 nf49e09.s  
c 767 27 AQ244398  
c 768 27 AQ526160 HS\_5312\_B  
c 769 27 AQ512553 HS\_5119\_B  
c 770 27 AQ562887  
c 771 27 AI732085  
c 772 27 AQ512745  
c 773 27 BE178481  
c 774 27 BE065970 RC3-HT060  
c 775 27 AW167073  
c 776 27 BF737749  
c 777 27 BI039053  
c 778 27 BI039054  
c 779 27 BG009072  
c 780 27 AU144449  
c 781 27 AV720925  
c 782 27 AQ536676 RPCI-11-3  
c 783 27 AQ691213 HS\_5338\_A  
c 784 27 BG056397 na6f68b04  
c 785 27 AQ078350 HS\_5569\_A  
c 786 27 AQ08037  
c 787 27 AL042731  
c 788 27 AL042731  
c 789 27 BI021436  
c 790 27 AQ480181  
c 791 27 AQ332193  
c 792 27 BI039969  
c 793 27 BF448499  
c 794 27 AW975367  
c 795 27 H23506  
c 796 27 AZ254674  
c 797 27 BF826816  
c 798 27 AW939214  
c 799 27 AQ570204  
c 800 27 AQ721483  
c 801 27 AQ352086  
c 802 27 AQ000610  
c 803 27 AQ020745  
c 804 27 BE740827  
c 805 27 BG983313  
c 806 27 N78063  
c 807 27 AQ678716  
c 808 27 BE704103  
c 809 27 AQ334955  
c 810 27 AQ533504  
c 811 27 BF675599  
c 812 27 BE143857  
c 813 27 BI0354690  
c 814 27 BF954499  
c 815 27 BE299730  
c 816 27 AQ441794 HS\_5106\_B  
c 817 27 AW603818 CN0-CN004  
c 818 27 AL044793 DKF2p434N  
c 819 27 AW471092 xu06h04.x  
c 820 27 AI500640 nt93e02.x  
c 821 27 AQ422506 RPCI-11-1  
c 822 27 BG777252 602664483  
c 823 27 BF919931 QVO-NT015  
c 824 27 B81834  
c 825 27 B81834  
c 826 27 AQ496808  
c 827 27 AQ620271  
c 828 27 AQ620271 HS\_5182\_B  
c 829 27 AQ226133 HS\_2015\_B  
c 830 27 AQ489571 RPCI-11-2  
c 831 27 AL597533 DKF2p313G  
c 832 27 AV739938 AV739938  
c 833 27 AA626478 ab53h12.f  
c 834 27 AQ429861 HS\_5077\_A  
c 835 27 AV739810 AV739810  
c 836 27 AA492114 ng17d03.s  
c 837 27 AA400544 zu62b12.f  
c 838 27 AQ211741 HS\_3239\_A  
c 839 27 AA670456 ae62h06.s  
c 840 27 AA482950 nf49e09.s  
c 841 27 AQ244398  
c 842 27 AQ526160 HS\_5312\_B  
c 843 27 AQ512553 HS\_5119\_B  
c 844 27 AQ562887  
c 845 27 AI732085  
c 846 27 AQ512745  
c 847 27 BE178481  
c 848 27 BE065970 RC3-HT060  
c 849 27 AW167073  
c 850 27 BF737749  
c 851 27 BI039053  
c 852 27 BI039054  
c 853 27 BG009072  
c 854 27 AU144449  
c 855 27 AV720925  
c 856 27 AQ536676 RPCI-11-3  
c 857 27 AQ691213 HS\_5338\_A  
c 858 27 BG056397 na6f68b04  
c 859 27 AQ078350 HS\_5569\_A  
c 860 27 AQ08037  
c 861 27 AL042731  
c 862 27 AL042731  
c 863 27 BI021436  
c 864 27 AQ480181  
c 865 27 AQ332193  
c 866 27 BI039969  
c 867 27 BF448499  
c 868 27 AW975367  
c 869 27 H23506  
c 870 27 AZ254674  
c 871 27 BF826816  
c 872 27 AW939214  
c 873 27 AQ570204  
c 874 27 AQ721483  
c 875 27 AQ352086  
c 876 27 AQ000610  
c 877 27 AQ020745  
c 878 27 BE740827  
c 879 27 BG983313  
c 880 27 N78063  
c 881 27 AQ678716  
c 882 27 BE704103  
c 883 27 AQ334955  
c 884 27 AQ533504  
c 885 27 BF675599  
c 886 27 BE143857  
c 887 27 BI0354690  
c 888 27 BF954499  
c 889 27 BE299730  
c 890 27 AQ441794 HS\_5106\_B  
c 891 27 AW603818 CN0-CN004  
c 892 27 AL044793 DKF2p434N  
c 893 27 AW471092 xu06h04.x  
c 894 27 AI500640 nt93e02.x  
c 895 27 AQ422506 RPCI-11-1  
c 896 27 BG777252 602664483  
c 897 27 BF919931 QVO-NT015  
c 898 27 B81834  
c 899 27 B81834  
c 900 27 AQ496808  
c 901 27 AQ620271  
c 902 27 AQ620271 HS\_5182\_B  
c 903 27 AQ226133 HS\_2015\_B  
c 904 27 AQ489571 RPCI-11-2  
c 905 27 AL597533 DKF2p313G  
c 906 27 AV739938 AV739938  
c 907 27 AA626478 ab53h12.f  
c 908 27 AQ429861 HS\_5077\_A  
c 909 27 AV739810 AV739810  
c 910 27 AA492114 ng17d03.s  
c 911 27 AA400544 zu62b12.f  
c 912 27 AQ211741 HS\_3239\_A  
c 913 27 AA670456 ae62h06.s  
c 914 27 AA482950 nf49e09.s  
c 915 27 AQ244398  
c 916 27 AQ526160 HS\_5312\_B  
c 917 27 AQ512553 HS\_5119\_B  
c 918 27 AQ562887  
c 919 27 AI732085  
c 920 27 AQ512745  
c 921 27 BE178481  
c 922 27 BE065970 RC3-HT060  
c 923 27 AW167073  
c 924 27 BF737749  
c 925 27 BI039053  
c 926 27 BI039054  
c 927 27 BG009072  
c 928 27 AU144449  
c 929 27 AV720925  
c 929 27 AQ536676 RPCI-11-3  
c 930 27 AQ691213 HS\_5338\_A  
c 931 27 BG056397 na6f68b04  
c 932 27 AQ078350 HS\_5569\_A  
c 933 27 AQ08037  
c 934 27 AL042731  
c 935 27 AL042731  
c 936 27 BI021436  
c 937 27 AQ480181  
c 938 27 AQ332193  
c 939 27 BI039969  
c 940 27 BF448499  
c 941 27 AW975367  
c 942 27 H23506  
c 943 27 AZ254674  
c 944 27 BF826816  
c 945 27 AW939214  
c 946 27 AQ570204  
c 947 27 AQ721483  
c 948 27 AQ352086  
c 949 27 AQ000610  
c 950 27 AQ020745  
c 951 27 BE740827  
c 952 27 BG983313  
c 953 27 N78063  
c 954 27 AQ678716  
c 955 27 BE704103  
c 956 27 AQ334955  
c 957 27 AQ533504  
c 958 27 BF675599  
c 959 27 BE143857  
c 960 27 BI0354690  
c 961 27 BF954499  
c 962 27 BE299730  
c 963 27 AQ441794 HS\_5106\_B  
c 964 27 AW603818 CN0-CN004  
c 965 27 AL044793 DKF2p434N  
c 966 27 AW471092 xu06h04.x  
c 967 27 AI500640 nt93e02.x  
c 968 27 AQ422506 RPCI-11-1  
c 969 27 BG777252 602664483  
c 970 27 BF919931 QVO-NT015  
c 971 27 B81834  
c 972 27 B81834  
c 973 27 AQ496808  
c 974 27 AQ620271  
c 975 27 AQ620271 HS\_5182\_B  
c 976 27 AQ226133 HS\_2015\_B  
c 977 27 AQ489571 RPCI-11-2  
c 978 27 AL597533 DKF2p313G  
c 979 27 AV739938 AV739938  
c 980 27 AA626478 ab53h12.f  
c 981 27 AQ429861 HS\_5077\_A  
c 982 27 AV739810 AV739810  
c 983 27 AA492114 ng17d03.s  
c 984 27 AA400544 zu62b12.f  
c 985 27 AQ211741 HS\_3239\_A  
c 986 27 AA670456 ae62h06.s  
c 987 27 AA482950 nf49e09.s  
c 988 27 AQ244398  
c 989 27 AQ526160 HS\_5312\_B  
c 990 27 AQ512553 HS\_5119\_B  
c 991 27 AQ562887  
c 992 27 AI732085  
c 993 27 AQ512745  
c 994 27 BE178481  
c 995 27 BE065970 RC3-HT060  
c 996 27 AW167073  
c 997 27 BF737749  
c 998 27 BI039053  
c 999 27 BI039054  
c 1000 27 BG009072

27 2.6 561 13 AQ562764  
27 2.6 562 10 AW889786  
27 2.6 563 10 AU147926  
27 2.6 564 10 AQ568574  
27 2.6 565 13 AU147352  
27 2.6 566 11 N25784  
27 2.6 567 13 AQ791303  
27 2.6 568 13 AW409893  
27 2.6 569 13 AQ547583  
27 2.6 570 13 AQ346926  
27 2.6 571 11 BE884181  
27 2.6 572 13 AQ547583  
27 2.6 573 13 AQ346926  
27 2.6 574 11 BE884181  
27 2.6 575 10 AW965632  
27 2.6 576 10 AW975952  
27 2.6 577 10 AW965632  
27 2.6 578 10 AL046677  
27 2.6 579 11 BF941382  
27 2.6 580 10 AV718803  
27 2.6 581 13 B53056  
27 2.6 582 10 AA428512  
27 2.6 583 13 AQ636023  
27 2.6 584 11 N58473  
27 2.6 585 10 AL577347  
27 2.6 586 13 AQ321791  
27 2.6 587 10 BE160626  
27 2.6 588 10 BE160626  
27 2.6 589 10 AV723671  
27 2.6 590 10 BE143624  
27 2.6 591 13 AG019176  
27 2.6 592 11 BF905405  
27 2.6 593 10 BF905405  
27 2.6 594 10 AQ383938  
27 2.6 595 10 AA155784  
27 2.6 596 11 W81268  
27 2.6 597 11 BG570417  
27 2.6 598 10 AW603838  
27 2.6 599 10 AA133332  
27 2.6 600 11 BG506930  
27 2.6 601 13 AQ666530  
27 2.6 602 10 AQ939447  
27 2.6 603 10 AW853213  
27 2.6 604 10 AW702975  
27 2.6 605 13 AQ040460  
27 2.6 606 13 AQ279584  
27 2.6 607 13 AQ312956  
27 2.6 608 13 AQ025882  
27 2.6 609 10 BE147973  
27 2.6 610 13 AQ267293  
27 2.6 611 13 AG025922  
27 2.6 612 10 A1133514  
27 2.6 613 11 BG392409  
27 2.6 614 11 BG542392  
27 2.6 615 10 AW962959  
27 2.6 616 13 AQ039461  
27 2.6 617 13 AQ314145  
27 2.6 618 10 AV758902  
27 2.6 619 13 AG019084  
27 2.6 620 11 BG202532  
27 2.6 621 10 AW969652  
27 2.6 622 10 AV763950  
27 2.6 623 10 BE273856  
27 2.6 624 13 AQ528747  
27 2.6 625 13 AQ037792  
27 2.6 626 13 AQ630943  
27 2.6 627 10 AV760878  
27 2.6 628 10 AV704876  
27 2.6 629 10 BE545198  
27 2.6 630 13 AQ351773  
27 2.6 631 11 BG536089  
27 2.6 632 13 AQ543328  
27 2.6 633 13 AQ562764  
27 2.6 634 10 AW889786  
27 2.6 635 10 AU147926  
27 2.6 636 13 AU147352  
27 2.6 637 11 N25784  
27 2.6 638 13 AQ791303  
27 2.6 639 13 AW409893  
27 2.6 640 13 AQ547583  
27 2.6 641 13 AQ346926  
27 2.6 642 11 BE884181  
27 2.6 643 10 AW965632  
27 2.6 644 10 AW975952  
27 2.6 645 10 AW965632  
27 2.6 646 10 AL046677  
27 2.6 647 11 BF941382  
27 2.6 648 10 AV718803  
27 2.6 649 13 B53056  
27 2.6 650 10 AA428512  
27 2.6 651 13 AQ636023  
27 2.6 652 11 N58473  
27 2.6 653 10 AL577347  
27 2.6 654 13 AQ321791  
27 2.6 655 10 BE160626  
27 2.6 656 10 BE160626  
27 2.6 657 10 AV723671  
27 2.6 658 10 BE143624  
27 2.6 659 13 AG019176  
27 2.6 660 11 BF905405  
27 2.6 661 10 BF905405  
27 2.6 662 10 AQ383938  
27 2.6 663 10 AA155784  
27 2.6 664 11 W81268  
27 2.6 665 11 BG570417  
27 2.6 666 10 AW603838  
27 2.6 667 10 AA133332  
27 2.6 668 11 BG506930  
27 2.6 669 13 AQ666530  
27 2.6 670 10 AQ939447  
27 2.6 671 10 AW853213  
27 2.6 672 10 AW702975  
27 2.6 673 13 AQ040460  
27 2.6 674 13 AQ279584  
27 2.6 675 13 AQ312956  
27 2.6 676 13 AQ025882  
27 2.6 677 10 BE147973  
27 2.6 678 13 AQ267293  
27 2.6 679 13 AG025922  
27 2.6 680 10 A1133514  
27 2.6 681 11 BG392409  
27 2.6 682 11 BG542392  
27 2.6 683 10 AW962959  
27 2.6 684 13 AQ039461  
27 2.6 685 13 AQ314145  
27 2.6 686 10 AV758902  
27 2.6 687 13 AG019084  
27 2.6 688 11 BG202532  
27 2.6 689 10 AW969652  
27 2.6 690 10 AV763950  
27 2.6 691 10 BE273856  
27 2.6 692 13 AQ528747  
27 2.6 693 13 AQ037792  
27 2.6 694 13 AQ630943  
27 2.6 695 10 AV760878  
27 2.6 696 10 AV704876  
27 2.6 697 10 BE545198  
27 2.6 698 13 AQ351773  
27 2.6 699 11 BG536089  
27 2.6 700 13 AQ543328  
27 2.6 701 13 AQ562764  
27 2.6 702 10 AW889786  
27 2.6 703 10 AU147926  
27 2.6 704 13 AU147352  
27 2.6 705 11 N25784  
27 2.6 706 13 AQ791303  
27 2.6 707 13 AW409893  
27 2.6 708 13 AQ547583  
27 2.6 709 13 AQ346926  
27 2.6 710 11 BE884181  
27 2.6 711 10 AW965632  
27 2.6 712 10 AW975952  
27 2.6 713 10 AW965632  
27 2.6 714 10 AL046677  
27 2.6 715 11 BF941382  
27 2.6 716 10 AV718803  
27 2.6 717 13 B53056  
27 2.6 718 10 AA428512  
27 2.6 719 13 AQ636023  
27 2.6 720 11 N58473  
27 2.6 721 10 AL577347  
27 2.6 722 13 AQ321791  
27 2.6 723 10 BE160626  
27 2.6 724 10 BE160626  
27 2.6 725 10 AV723671  
27 2.6 726 10 BE143624  
27 2.6 727 13 AG019176  
27 2.6 728 11 BF905405  
27 2.6 729 10 BF905405  
27 2.6 730 10 AQ383938  
27 2.6 731 10 AA155784  
27 2.6 732 11 W81268  
27 2.6 733 11 BG570417  
27 2.6 734 10 AW603838  
27 2.6 735 10 AA133332  
27 2.6 736 11 BG506930  
27 2.6 737 13 AQ666530  
27 2.6 738 10 AQ939447  
27 2.6 739 10 AW853213  
27 2.6 740 10 AW702975  
27 2.6 741 13 AQ040460  
27 2.6 742 13 AQ279584  
27 2.6 743 13 AQ312956  
27 2.6 744 13 AQ025882  
27 2.6 745 10 BE147973  
27 2.6 746 13 AQ267293  
27 2.6 747 13 AG025922  
27 2.6 748 10 A1133514  
27 2.6 749 11 BG392409  
27 2.6 750 11 BG542392  
27 2.6 751 10 AW962959  
27 2.6 752 13 AQ039461  
27 2.6 753 13 AQ314145  
27 2.6 754 10 AV758902  
27 2.6 755 13 AG019084  
27 2.6 756 11 BG202532  
27 2.6 757 10 AW969652  
27 2.6 758 10 AV763950  
27 2.6 759 10 BE273856  
27 2.6 760 13 AQ528747  
27 2.6 761 13 AQ037792  
27 2.6 762 13 AQ630943  
27 2.6 763 10 AV760878  
27 2.6 764 10 AV704876  
27 2.6 765 10 BE545198  
27 2.6 766 13 AQ351773  
27 2.6 767 11 BG536089  
27 2.6 768 13 AQ543328  
27 2.6 769 13 AQ562764  
27 2.6 770 10 AW889786  
27 2.6 771 10 AU147926  
27 2.6 772 13 AU147352  
27 2.6 773 11 N25784  
27 2.6 774 13 AQ791303  
27 2.6 775 13 AW409893  
27 2.6 776 13 AQ547583  
27 2.6 777 13 AQ346926  
27 2.6 778 11 BE884181  
27 2.6 779 10 AW965632  
27 2.6 780 10 AW975952  
27 2.6 781 10 AW965632  
27 2.6 782 10 AL046677  
27 2.6 783 11 BF941382  
27 2.6 784 10 AV718803  
27 2.6 785 13 B53056  
27 2.6 786 10 AA428512  
27 2.6 787 13 AQ636023  
27 2.6 788 11 N58473  
27 2.6 789 10 AL577347  
27 2.6 790 13 AQ321791  
27 2.6 791 10 BE160626  
27 2.6 792 10 BE160626  
27 2.6 793 10 AV723671  
27 2.6 794 10 BE143624  
27 2.6 795 13 AG019176  
27 2.6 796 11 BF905405  
27 2.6 797 10 BF905405  
27 2.6 798 10 AQ383938  
27 2.6 799 10 AA155784  
27 2.6 800 11 W81268  
27 2.6 801 11 BG570417  
27 2.6 802 10 AW603838  
27 2.6 803 10 AA133332  
27 2.6 804 11 BG506930  
27 2.6 805 13 AQ666530  
27 2.6 806 10 AQ939447  
27 2.6 807 10 AW853213  
27 2.6 808 10 AW702975  
27 2.6 809 13 AQ040460  
27 2.6 810 13 AQ279584  
27 2.6 811 13 AQ312956  
27 2.6 812 13 AQ025882  
27 2.6 813 10 BE147973  
27 2.6 814 13 AQ267293  
27 2.6 815 13 AG025922  
27 2.6 816 10 A1133514  
27 2.6 817 11 BG392409  
27 2.6 818 11 BG542392  
27 2

C 816	27	2.6	696	10	AV718208	AV718208	AV718208	AV718208	889	27	2.6	1071	10	AL569845	AL569845
C 817	27	2.6	696	13	AG016448	AG016448	Homo sapi	AG016448	C 890	27	2.6	1079	11	BG698721	BG698721
C 818	27	2.6	697	10	AV762945	AV762945	AV762945	AV762945	C 891	27	2.6	1091	11	BF681181	BF681181
C 819	27	2.6	697	13	AG016444	AG016444	Homo sapi	AG016444	C 892	27	2.6	1129	11	BG775312	BG775312
C 820	27	2.6	697	13	AG016444	AG016444	Homo sapi	AG016444	C 893	27	2.6	1129	12	AF305824	AF305824
C 821	27	2.6	698	10	AU117514	AU117514	AU117514	AU117514	C 894	27	2.6	1187	11	BG251369	BG251369
C 822	27	2.6	699	10	BE169990	BE169990	CMO-HT053	BE169990	C 895	27	2.6	1243	10	AV763002	AV763002
C 823	27	2.6	699	13	AG017056	AG017056	AG017056	AG017056	C 896	27	2.6	1247	10	AV761200	AV761200
C 824	27	2.6	699	13	AG013234	AG013234	Homo sapi	AG013234	C 897	27	2.6	1335	10	AV761381	AV761381
C 825	27	2.6	704	13	AG016449	AG016449	Homo sapi	AG016449	C 898	27	2.6	1474	10	AV762221	AV762221
C 826	27	2.6	705	10	AV707907	AV707907	AV707907	AV707907	C 899	27	2.6	1639	12	AF113684	AF113684
C 827	27	2.6	709	13	AG013216	AG013216	Homo sapi	AG013216	C 900	27	2.6	1898	12	AF113684	AF113684
C 828	27	2.6	711	11	BF666079	BF666079	BF666079	BF666079	C 901	26	2.5	168	13	AA288137	AA288137
C 829	27	2.6	714	10	AV713462	AV713462	AV713462	AV713462	C 902	26	2.5	169	13	B55051	B55051
C 830	27	2.6	716	13	AG013215	AG013215	Homo sapi	AG013215	C 903	26	2.5	182	13	AA795338	AA795338
C 831	27	2.6	718	10	AW963636	AW963636	BF665666	BF665666	C 904	26	2.5	194	10	AA658182	AA658182
C 832	27	2.6	720	11	BF968561	BF968561	BF968561	BF968561	C 905	26	2.5	197	11	T51743	T51743
C 833	27	2.6	723	11	BG502627	BG502627	BG502627	BG502627	C 906	26	2.5	201	11	BF886789	BF886789
C 834	27	2.6	729	10	AV762685	AV762685	AV762685	AV762685	C 907	26	2.5	205	10	AW081911	AW081911
C 835	27	2.6	730	10	AV760533	AV760533	AV760533	AV760533	C 908	26	2.5	214	11	BF991590	BF991590
C 836	27	2.6	730	11	BG529467	BG529467	BG529467	BG529467	C 909	26	2.5	216	11	R24500	R24500
C 837	27	2.6	732	11	BF666566	BF666566	BF666566	BF666566	C 910	26	2.5	218	11	T12423	T12423
C 838	27	2.6	732	11	BG289613	BG289613	BG289613	BG289613	C 911	26	2.5	218	13	B71710	B71710
C 839	27	2.6	733	11	BG675576	BG675576	BG675576	BG675576	C 912	26	2.5	221	11	N70241	N70241
C 840	27	2.6	734	11	BG534528	BG534528	BG534528	BG534528	C 913	26	2.5	226	10	AA344084	AA344084
C 841	27	2.6	736	11	BG179169	BG179169	BG179169	BG179169	C 914	26	2.5	226	10	BE092802	BE092802
C 842	27	2.6	737	10	AL601138	AL601138	AL601138	AL601138	C 915	26	2.5	233	10	AW905647	AW905647
C 843	27	2.6	742	10	AV762133	AV762133	AV762133	AV762133	C 916	26	2.5	233	11	BF813493	BF813493
C 844	27	2.6	742	13	AG011957	AG011957	Homo sapi	AG011957	C 917	26	2.5	234	11	F34457	F34457
C 845	27	2.6	745	13	AQ895263	AQ895263	HS_3148_B	AQ895263	C 918	26	2.5	262	10	AI678542	AI678542
C 846	27	2.6	751	11	B1084691	B1084691	B1084691	B1084691	C 919	26	2.5	263	10	AW937369	AW937369
C 847	27	2.6	752	10	AV758960	AV758960	AV758960	AV758960	C 920	26	2.5	266	10	AA827295	AA827295
C 848	27	2.6	753	10	BE299494	BE299494	BE299494	BE299494	C 921	26	2.5	266	13	AO677178	AO677178
C 849	27	2.6	753	11	BF671285	BF671285	BF671285	BF671285	C 922	26	2.5	266	13	AO677178	AO677178
C 850	27	2.6	755	13	AG011956	AG011956	Homo sapi	AG011956	C 923	26	2.5	270	10	AA303165	AA303165
C 851	27	2.6	761	10	AV761908	AV761908	AV761908	AV761908	C 924	26	2.5	272	10	AA303165	AA303165
C 852	27	2.6	761	11	BG681910	BG681910	BG681910	BG681910	C 925	26	2.5	273	11	BF914788	BF914788
C 853	27	2.6	765	11	BF036900	BF036900	BF036900	BF036900	C 926	26	2.5	275	10	AA635067	AA635067
C 854	27	2.6	765	13	AQ896662	AQ896662	HS_3117_B	AQ896662	C 927	26	2.5	295	10	AA376615	AA376615
C 855	27	2.6	769	10	AL521028	AL521028	AL521028	AL521028	C 928	26	2.5	302	10	AA838096	AA838096
C 856	27	2.6	770	10	AL582702	AL582702	AL582702	AL582702	C 929	26	2.5	302	10	AW934751	AW934751
C 857	27	2.6	782	11	BG434786	BG434786	BG434786	BG434786	C 930	26	2.5	305	11	BG997683	BG997683
C 858	27	2.6	794	11	BG740010	BG740010	BG740010	BG740010	C 931	26	2.5	312	10	AW805169	AW805169
C 859	27	2.6	829	11	BG911559	BG911559	BG911559	BG911559	C 932	26	2.5	314	10	AW085785	AW085785
C 860	27	2.6	829	11	BF692882	BF692882	BF692882	BF692882	C 933	26	2.5	316	13	B78760	B78760
C 861	27	2.6	837	11	BF692882	BF692882	BF692882	BF692882	C 934	26	2.5	319	10	AW837332	AW837332
C 862	27	2.6	842	13	B02326	B02326	PT2_1_13-	B02326	C 935	26	2.5	326	11	BE139033	BE139033
C 863	27	2.6	845	10	AI557106	AI557106	AI557106	AI557106	C 936	26	2.5	326	11	BI062186	BI062186
C 864	27	2.6	851	11	BG283407	BG283407	BG283407	BG283407	C 937	26	2.5	327	10	AA090479	AA090479
C 865	27	2.6	858	11	BF214379	BF214379	BF214379	BF214379	C 938	26	2.5	337	10	AA090479	AA090479
C 866	27	2.6	859	11	BG615899	BG615899	BG615899	BG615899	C 939	26	2.5	344	10	AW920266	AW920266
C 867	27	2.6	891	11	BE972613	BE972613	BE972613	BE972613	C 940	26	2.5	346	10	AA729166	AA729166
C 868	27	2.6	906	10	BE784083	BE784083	BE784083	BE784083	C 941	26	2.5	347	10	AA847357	AA847357
C 869	27	2.6	914	11	BG912185	BG912185	BG912185	BG912185	C 942	26	2.5	352	11	R89367	R89367
C 870	27	2.6	917	11	BE885138	BE885138	BE885138	BE885138	C 943	26	2.5	354	10	AW817527	AW817527
C 871	27	2.6	924	11	BG388300	BG388300	BG388300	BG388300	C 944	26	2.5	355	10	AI420666	AI420666
C 872	27	2.6	924	11	BG697538	BG697538	BG697538	BG697538	C 945	26	2.5	356	13	BI5838	BI5838
C 873	27	2.6	935	11	BG432843	BG432843	BG432843	BG432843	C 946	26	2.5	357	13	AO665282	AO665282
C 874	27	2.6	935	11	BG677896	BG677896	BG677896	BG677896	C 947	26	2.5	358	11	BF843520	BF843520
C 875	27	2.6	943	10	AL527924	AL527924	AL527924	AL527924	C 948	26	2.5	360	10	AA468990	AA468990
C 876	27	2.6	946	10	AL570329	AL570329	AL570329	AL570329	C 949	26	2.5	364	10	AW861716	AW861716
C 877	27	2.6	956	11	BG110329	BG110329	BG110329	BG110329	C 950	26	2.5	366	10	AA501867	AA501867
C 878	27	2.6	969	11	BG616087	BG616087	BG616087	BG616087	C 951	26	2.5	366	10	AV762966	AV762966
C 879	27	2.6	970	10	BE788445	BE788445	BE788445	BE788445	C 952	26	2.5	368	11	H82863	H82863
C 880	27	2.6	976	10	AL553425	AL553425	AL553425	AL553425	C 953	26	2.5	369	11	AA618172	AA618172
C 881	27	2.6	978	11	BF027503	BF027503	BF027503	BF027503	C 954	26	2.5	369	11	T51589	T51589
C 882	27	2.6	989	11	BG116144	BG116144	BG116144	BG116144	C 955	26	2.5	372	13	AO278367	AO278367
C 883	27	2.6	990	10	AL525925	AL525925	AL525925	AL525925	C 956	26	2.5	373	13	AO278367	AO278367
C 884	27	2.6	1005	10	AL566411	AL566411	AL566411	AL566411	C 957	26	2.5	375	10	AA507474	AA507474
C 885	27	2.6	1005	11	BG706422	BG706422	BG706422	BG706422	C 958	26	2.5	376	10	AW827426	AW827426
C 886	27	2.6	1024	11	BG706427	BG706427	BG706427	BG706427	C 959	26	2.5	377	13	AO631460	AO631460
C 887	27	2.6	1027	11	BG622059	BG622059	BG622059	BG622059	C 960	26	2.5	383	11	BG152452	BG152452
C 888	27	2.6	1036	10	BE299695	BE299695	BE299695	BE299695	C 961	26	2.5	394	11	BI034667	BI034667

```
c 962 26 2.5 394 11 BF963551
963 26 2.5 399 10 AA679235
964 26 2.5 399 10 AA679235
965 26 2.5 399 10 AA679235
966 26 2.5 401 10 AV738723
967 26 2.5 403 10 AA973261
968 26 2.5 403 10 AA973261
969 26 2.5 404 13 AQ009396
970 26 2.5 406 13 AQ009396
971 26 2.5 406 13 AQ009396
972 26 2.5 408 11 BF872591
973 26 2.5 411 11 H53579
974 26 2.5 412 10 AA835938
975 26 2.5 412 10 AA835938
976 26 2.5 414 11 BF943816
977 26 2.5 415 11 T95842
978 26 2.5 418 10 AA081047
979 26 2.5 418 10 AA310218
980 26 2.5 418 11 BF901094
981 26 2.5 419 10 AA152074
982 26 2.5 420 13 B99694
983 26 2.5 421 13 AQ060139
984 26 2.5 423 10 AA113126
985 26 2.5 425 10 AA436716
986 26 2.5 425 13 AZ520178
987 26 2.5 427 10 AA121862
988 26 2.5 427 10 AA121862
989 26 2.5 427 11 BF954832
990 26 2.5 428 10 AA632895
991 26 2.5 429 11 BF964095
992 26 2.5 430 10 AA174110
993 26 2.5 430 11 R27323
994 26 2.5 430 11 R27323
995 26 2.5 431 10 AA055882
996 26 2.5 433 13 AQ114206
997 26 2.5 434 10 AA811150
998 26 2.5 435 10 AA969798
999 26 2.5 436 10 AA649183
1000 26 2.5 436 10 AA775201
```

## ALIGNMENTS

```
RESULT 1
LOCUS AA626859 295 bp mRNA EST 15-OCT-1997
DEFINITION zu89f10.s1 Soares, testis_NHT Homo sapiens cDNA clone IMAGE:745195
3', similar to SW:KKIA_HUMAN Q00532 SERINE/THREONINE-PROTEIN KINASE
KIALRE.; mRNA sequence.
ACCESSION AA626859
VERSION AA626859.1 GI:2539246
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 295)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisinger,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham
```

```
FEATURES
source High quality sequence stop: 292.
Location/Qualifiers
1..295
/organism="Homo sapiens"
/db_xref="GDB:5933055"
/db_xref="Taxon:9606"
/clone="IMAGE:745195"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATGTAAGTGGAGCGGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 95 a 69 c 55 g 76 t
ORIGIN
Query Match 19.0%; Score 198; DB 10; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.5e-90;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 455 ttccagagagatcaccacattatgtagctacgagatgctacgagctcgaacttc 514
Db 263 TTCCAGAGATGCTTACACCAATATGTAGCTACGAGATGTCACGAGCTCTGACATC 204
QY 515 ttgtgggaatactcagtagttcttcagtcgatatagggctattggtgtgtttg 574
Db 203 TTGTGGAGATATCTCAGTATGTTCTTCAGTCGATATATGGCTATTTGTTTGG 144
QY 575 cagagctctgcagccagccagctgctgcagaaatcagatgggacacatttatc 634
Db 143 CAGAGCTCTGCAGAGCCAGCCACTGTGGCTTGGAAATCAGATGTGGACCACTTATC 84
QY 635 tgataatcagaacactag 652
Db 83 TGATAATCAGAACACTAG 66
RESULT 2
BG390099 756 bp mRNA EST 12-MAR-2001
LOCUS 602415940F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4524513 5',
DEFINITION mRNA sequence.
ACCESSION BG390099
VERSION BG390099.1 GI:13283645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 756)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10428 row: o column: 10
High quality sequence stop: 748.
Location/Qualifiers
FEATURES
```

```
source
1..756
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/notes="Organ: Testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 205 a 177 c 192 g 182 t
ORIGIN

Query Match 4.5%; Score 47; DB 11; Length 756;
Best Local Similarity 100.0%; Pred. No. 5e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagcgggttcgcatgtgacacgagcgtgctcgaactc 696
|||||
Db 7 TAGTAGAGAGCGGGTTTCGCCATGTTGACCGAGCTGTCTCGAATC 53

RESULT 3
BF869052/c
LOCUS BF869052 433 bp mRNA EST 17-JAN-2001
DEFINITION IL5-ET0119-181000-180-f09 ET0119 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF869052
VERSION BF869052.1 GI:12259182
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 433)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5et2-IL5-ET0119-
181000-180-f09at3-2000-10-18at4-1)
Seq primer: puc 18 forward
High quality sequence stop: 396.
Location/Qualifiers
1..433
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0119"
/dev_stage="Adult"
/notes="Organ: Lung,tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under
```

```
low stringency conditions."
BASE COUNT 137 a 102 c 111 g 83 t
ORIGIN

Query Match 3.9%; Score 41; DB 11; Length 433;
Best Local Similarity 100.0%; Pred. No. 6.2e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagcgggttcgcatgtgacacgagcgtgctcgc 690
|||||
Db 124 TAGTAGAGAGCGGGTTTCGCCATGTTGACCGAGCTGTCTC 84

RESULT 4
AQ230677
LOCUS AQ230677 328 bp DNA GSS 26-SEP-1998
DEFINITION HS_2033_B2_A10_MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-2033 Col=20 Row=B, DNA sequence.
ACCESSION AQ230677
VERSION AQ230677.1 GI:3655906
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 328)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2033 row: B column: 20
Class: BAC ends
High quality sequence stop: 328.
Location/Qualifiers
1..328
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Plate-2033 Col=20 Row=B"
/sex="male"
/notes="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 67 a 94 c 69 g 98 t
ORIGIN

Query Match 3.7%; Score 39; DB 13; Length 328;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagcgggttcgcatgtgacacgagcgtgctc 688
|||||
Db 232 TAGTAGAGAGCGGGTTTCGCCATGTTGACCGAGCTGTCTC 270

RESULT 5
BI040145
LOCUS BI040145 171 bp mRNA EST 14-JUN-2001
DEFINITION RC5-NT0263-010201-012-F07 NT0263 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI040145
VERSION BI040145.1 GI:14446771
```

KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 171)

DIAS NETO, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.J.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&t2=RC5-NT0263-010201-012-F07&t3=2001-02-01&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence stop: 171.

FEATURES  
source Location/Qualifiers  
1..171  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NT0263"  
/dev\_stage="Adult"  
/note="Organ: nervous tumor; Vector: puc18; Site: 1: Smal; Site 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
39 a 43 c 45 g 44 t  
BASE COUNT 39 a 43 c 45 g 44 t  
ORIGIN

Query Match 3.6%; Score 37; DB 11; Length 171;  
Best Local Similarity 100.0%; Pred. No. 7.2e-08;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 ggggttcgccatgtgaccaggctggtctgaactc 696  
|||||  
Db 23 ggggttcgccatgtgaccaggctggtctgaactc 59  
|||||

RESULT 6  
AQ030066  
LOCUS AQ030066 368 bp DNA GSS 14-APR-1999  
DEFINITION RPC111-39B18.TP RPC1-11 Homo sapiens genomic clone RPC1-11-39B18, DNA sequence.  
ACCESSION AQ030066  
VERSION AQ030066.1 GI:3274030  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 368)

REFERENCE  
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.  
TITLE Use of BAC End Sequences for Sequence-Ready Map Building (1998)

JOURNAL Unpublished (1998)  
COMMENT Other GSSs: RPC111-39B18.TV  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are derived from the human BAC library RPC1-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (<http://info@resgen.com>). BAC end search page: [http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html)  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source Location/Qualifiers  
1..368  
/organism="Homo sapiens"  
/db\_xref="GDB:7514633"  
/db\_xref="taxon:9606"  
/clone="RPC1-11-39B18"  
/clone\_lib="RPC1-11"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPC111 Human Male BAC Library"  
76 a 107 c 80 g 105 t  
BASE COUNT 76 a 107 c 80 g 105 t  
ORIGIN

Query Match 3.6%; Score 37; DB 13; Length 368;  
Best Local Similarity 100.0%; Pred. No. 7.1e-08;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 ggggttcgccatgtgaccaggctggtctgaactc 696  
|||||  
Db 70 ggggttcgccatgtgaccaggctggtctgaactc 106  
|||||

RESULT 7  
B32077  
LOCUS B32077 494 bp DNA GSS 17-OCT-1997  
DEFINITION HS-1014-B1-F09-MF abi CIT Human Genomic Sperm Library C Homo sapiens genomic clone Plate-CI 789 Col=17 Row=L, DNA sequence.  
ACCESSION B32077  
VERSION B32077.1 GI:2531446  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 494)

REFERENCE  
AUTHORS Mahairas, G.G., Zackrone, K.D., Smith, T., Tipton, S., Schmidt, S., Traicoff, R., Abajian, C., Blanchard, A., West, A., and Hood, L.E.  
TITLE Construction of a Characterized Clone Resource for Genomic Sequencing: Generation and Preliminary Analysis of 20,000 Sequence Tagged Connectors  
JOURNAL Unpublished (1997)  
COMMENT Contact: Mahairas GG; Zackrone KD, Hood L  
University of Washington  
Seattle, WA 98195, USA  
Tel: (206) 616-8744  
Fax: (206) 685-7301  
Email: kzackrone@u.washington.edu  
Sequence tagged Connector  
Plate: CT 789 row: L column: 17  
Class: BAC ends  
High quality sequence stop: 494.  
Location/Qualifiers  
1..494  
/organism="Homo sapiens"

us-09-671-050-9.oli20.rst

Fri May 3 11:02:12 2002

```

/db_xref="taxon:9606"
/clone="plate=CT 789 Col=17 Row=L"
/clone_lib="CIT Human Genomic Sperm Library C"
/sex="M"
/notes="Organ: sperm; Vector: pBelOAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      100 a 123 c 118 g 148 t 5 others
ORIGIN

Query Match      3.6%; Score 37; DB 13; Length 494;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 ggggttcgcatgttgaccaggctggtctgaactc 696
|||||
Db 320 GGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 356

RESULT 8
AG014845/c
LOCUS
DEFINITION
AG014845 AG006560
AG014845.1 GI:3650089
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE
Homo sapiens genomic DNA, chromosome 21q
JOURNAL
Published Only in Database (1998) In press
REFERENCE
2 (bases 1 to 700)
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE
Direct Submission
JOURNAL
Submitted (23-SEP-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@gc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
COMMENT
On Feb 6, 1999 this sequence version replaced gi:2992438.
AG006560: Submitted (27-Mar-1998).
FEATURES
source
1..700
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="289H18N11"
/map="21q"
BASE COUNT      134 a 152 c 156 g 137 t 121 others
ORIGIN

Query Match      3.6%; Score 37; DB 13; Length 700;
Best Local Similarity 100.0%; Pred. No. 6.9e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 ggggttcgcatgttgaccaggctggtctgaactc 696
|||||
Db 218 GGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 182

RESULT 9
BE150038
LOCUS
DEFINITION
QV3-HT0262-041199-014-b05 HT0262 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE150038
VERSION
BE150038.1 GI:8612759
KEYWORDS
EST.
SOURCE
human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 207)
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
COMMENT
Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV3-HT0262-041
199-014-b05&t3=1999-11-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 5
High quality sequence stop: 207.
Location/Qualifiers
1..207
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0262"
/dev_stage="Adult"
/notes="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      45 a 58 c 52 g 50 t 2 others
ORIGIN

Query Match      3.5%; Score 36; DB 10; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 gagacggggttcgccatgttgaccaggctggtctc 690
|||||
Db 67 GAGACGGGGTTCGCCATGTTGACCAGGCTGGTCTC 102

RESULT 10
H95588/c
LOCUS
DEFINITION
H95588 366 bp mRNA EST 25-NOV-1996
IMAGE:242905 3' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION
H95588
VERSION
H95588.1 GI:1108730
KEYWORDS
EST.
SOURCE
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 366)
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisakis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.

```





/clone="Plate=3155 Col=18 Row=D"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelOBAcl1; BAC Clones in E-Coli DH10B"  
BASE COUNT 143 a 110 c 131 g 110 t 6 others  
ORIGIN

Query Match 3.5%; Score 36; DB 13; Length 500;  
Best Local Similarity 100.0%; Pred. No. 2.3e-07;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 661 gggttcgcatgttgaccagctggtctgaactc 696  
|||||  
Db 330 GGGTTTCCCATGTTGACCAGGCTGGTCTGAACTC 295

RESULT 13  
LOCUS AQ668860/c 516 bp DNA GSS 24-JUN-1999  
DEFINITION HS\_5431\_Al\_E09\_SF6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1007 Col=17 Row=1, DNA sequence.  
ACCESSION AQ668860  
VERSION AQ668860.1 GI:5201606  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 516)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and Hood,L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
MEDLINE 99380589  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu  
Plate: 1007 row: 1 column: 17  
Seq primer: SP6  
Class: BAC ends  
High quality sequence stop: 516.  
Location/Qualifiers  
1..516  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clones="Plate=1007 Col=17 Row=1"  
/clone\_lib="RPCI-11 Human Male BAC Library"  
/sex="male"  
/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"  
BASE COUNT 164 a 107 c 103 g 137 t 5 others  
ORIGIN

Query Match 3.5%; Score 36; DB 13; Length 516;  
Best Local Similarity 100.0%; Pred. No. 2.3e-07;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 654 agagacgggtttcccatgttgaccaggtgctgtct 689  
|||||  
Db 339 ACACACGGGGTTTCGCCATGTTGACCAGGCTGTCT 304

RESULT 14  
LOCUS BG698963 801 bp mRNA 07-MAY-2001  
DEFINITION 602703384F1 NCL\_CGAP\_Skn3 Homo sapiens cDNA clone IMAGE:4800452 5', mRNA sequence.  
ACCESSION BG698963  
VERSION BG698963.1 GI:13966783  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 801)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: James Cleaver, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLAM10691 row: h column: 21  
High quality sequence stop: 768.  
Location/Qualifiers  
1..801  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4800452"  
/clone\_lib="NCL\_CGAP\_Skn3"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: skin; Vector: pCMV-SpORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCL\_CGAP Library."  
BASE COUNT 233 a 193 c 198 g 177 t  
ORIGIN

Query Match 3.5%; Score 36; DB 11; Length 801;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagagacgggtttcccatgttgaccaggtg 685  
|||||  
Db 264 TAGTAGACGGGGTTTCGCCATGTTGACCAGGCTG 229

RESULT 15  
LOCUS BE149958 193 bp mRNA 21-JUN-2000  
DEFINITION RC0-HT0259-201199-031-e09 HT0259 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE149958  
VERSION BE149958.1 GI:8612679  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 193)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,



(Clontech); Site\_1: sfii (ggccgctcgcc); Site\_2: sfii (ggccattatggc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 161 a 163 c 156 g 240 t  
ORIGIN

Query Match 3.4%; Score 35; DB 11; Length 720;  
Best Local Similarity 100.0%; Pred. NO. 7.3e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 662 gtttcgcacgttgaccaggctgtctcgaaactc 696  
Db 95 GGTTCGCCATGTTGACAGGCTGTCTCGAACTC 129

RESULT 18  
LOCUS AQ3211133 778 bp DNA GSS 06-MAY-1999  
DEFINITION RPCI11-111E3.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-111E3, DNA sequence.  
ACCESSION AQ3211133  
VERSION AQ3211133.1 GI:4053967  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1. (bases 1 to 778)  
Adams M.D., Rounsley S.D., Zhao S., Bass S., Linher K., Golden K., Berry K., Granger D., Suh E., Wible C., de Jong P. and Venter J.C. Use of human BAC End Sequences for Sequence-Ready Map Building Unpublished (1998)  
Other GSSs: RPCI11-111E3.TV  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source  
1. .778  
/organism="Homo sapiens"  
/db\_xref="GDB:7542338"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-111E3"  
/clone\_lib="RPCI-11"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPCI11 Human Male BAC Library"

BASE COUNT 176 a 185 c 146 g 271 t  
ORIGIN

Query Match 3.4%; Score 35; DB 13; Length 778;  
Best Local Similarity 100.0%; Pred. NO. 7.3e-07;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 650 tagtagacgagggtttccgcatgttgaccaggct 684  
Db 525 TAGTAGAGAGCGGGTTTCGCCATGTTGACCAGGCT 559

RESULT 19  
LOCUS BF878219 369 bp mRNA EST 17-JAN-2001  
DEFINITION CM2-ET0127-181100-523-f11 ET0127 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF878219  
VERSION BF878219.1 GI:12268349  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1. (bases 1 to 369)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and Simpson, A.J.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.J.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil.  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-ET0127-181100-523-f11&t3=2000-11-18&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 4  
High quality sequence stop: 369.  
Location/Qualifiers  
1. .369  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="ET0127"  
/dev\_stage="Adult"  
/note="Organ: lung\_tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 82 a 116 c 73 g 98 t  
ORIGIN

Query Match 3.3%; Score 34; DB 11; Length 369;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 663 gtttcgccatgttgaccaggctgtctcgaaactc 696  
Db 152 GTTTCGCCATGTTGACAGGCTGTCTCGAACTC 185

RESULT 20  
LOCUS AQ346239 377 bp DNA GSS 07-MAY-1999  
DEFINITION AQ346239

```

DEFINITION   RPCI11-116D16.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-116D16,
ACCESSION    AQ346239
VERSION      AQ346239.1 GI:4171135
KEYWORDS     GSS.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 377)
AUTHORS      Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
              ,J.C.
TITLE        Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
              Map Building
JOURNAL      Unpublished (1997)
COMMENT      Other GSSs: RPCI11-116D16.TV
              Contact: Shaying Zhao, William Nierman, Mark Adams
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: hsetigr.org
              Clones are derived from the human BAC library RPCI-11. For BAC
              library availability, please contact Pieter de Jong
              (pieter@dejong.med.buffalo.edu). Clones may be purchased from
              BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
              Research Genetics (http://bacpac.med.buffalo.edu/ordering) or from
              http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
              Seq primer: SP6
              Class: BAC ends.
FEATURES     source
              Location/Qualifiers
                1..377
                /organism="Homo sapiens"
                /db_xref="GDB:754247"
                /db_xref="taxon:9606"
                /clone="RPCI-11-116D16"
                /clone_lib="RPCI-11"
                /sex="Male"
                /cell_type="Lymphocytes"
                /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
              RPCI11 Human Male BAC Library"
BASE COUNT   77 a 114 c 79 g 107 t
ORIGIN

Query Match      3.3%; Score 34; DB 13; Length 377;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 663 gtttcgcatgttgaccaggctgtctcgaaactc 696
      |||||||
Db 84 GTTTCGCATGTTGACAGGCTGTCTCGAACTC 117

RESULT 21
AW864431
LOCUS      AW864431 407 bp mRNA EST 22-MAY-2000
DEFINITION PM4-SN0016-030400-002-b12 SN0016 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW864431
VERSION     AW864431.1 GI:7998390
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 407)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
              Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
              ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.

```

```

TITLE        Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202653
COMMENT      Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL
              (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-PM4-SN0016-030
              400-002-b12&t3=2000-04-03&t4=1)
              Seq primer: puc 18 forward
              High quality sequence stop: 382.
FEATURES     Source
              Location/Qualifiers
                1..407
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="SN0016"
                /dev_stage="Adult"
                /note="Organ: stomach.normal; Vector: puc18; Site_1: SmaI;
              Site_2: SmaI; A mini-library was made by cloning products
              derived from ORESTES PCR (U.S. Letters Patent application
              No. 196,716 - Ludwig Institute for Cancer Research)
              profiles into the pUC 18 vector. Reverse transcription of
              tissue mRNA and cDNA amplification were performed under
              low stringency conditions."
BASE COUNT   73 a 95 c 122 g 117 t
ORIGIN

Query Match      3.3%; Score 34; DB 10; Length 407;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 663 gtttcgcatgttgaccaggctgtctcgaaactc 696
      |||||||
Db 200 GTTTCGCATGTTGACAGGCTGTCTCGAACTC 233

RESULT 22
AQ093501/c
LOCUS      AQ093501 412 bp DNA GSS 27-AUG-1998
DEFINITION HS_3018_B1_G09_T7 CIT Approved Human Genomic Sperm Library D Homo
              sapiens genomic clone Plate=3018 Col=17 Row=N, DNA sequence.
ACCESSION  AQ093501
VERSION     AQ093501.1 GI:3464948
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 412)
AUTHORS      Maharras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
              Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
              Hood,L.
TITLE        Sequence-tagged connectors: A sequence approach to mapping and
              scanning the human genome
JOURNAL      Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE      99380589
COMMENT      Contact: Maharras GG, Wallace JC, Hood L
              High Throughput Sequencing Center
              University of Washington
              401 Queen Anne Avenue North, Seattle, WA 98109, USA
              Tel: (206) 616-3618
              Fax: (206) 616-3887
              Email: jwallace@u.washington.edu
              Sequence Tagged Connector
              Plate: 3018 row: N column: 17

```

Class: BAC ends  
High quality sequence stop: 412.  
Location/Qualifiers  
1..412  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate3018 Col-17 Row-N"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

140 a 94 c 95 g 83 t  
BASE COUNT  
ORIGIN

Query Match 3.3%; Score 34; DB 13; Length 412;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 663 gtttcgcatgttgaccaggctggtctcgaaactc 696  
|||||  
Db 210 GTTTCGCATGTTGACGAGCTGCTCGAACTC 177

RESULT 23  
AW958514  
LOCUS  
DEFINITION EST370584 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW958514  
VERSION AW958514.1 GI:8148198  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 589)  
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspar,R., Gay,C., Holt  
I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and  
Quackenbush,J.  
TITLE Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
UNPUBLISHED (2000)  
CONTACT: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnq@tigr.org  
Plate: 121  
Seq primer: Reverse.  
Location/Qualifiers  
1..589  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGE"  
/note="Vector: pBluescriptSKm"

153 a 131 c 131 g 174 t  
BASE COUNT  
ORIGIN

Query Match 3.3%; Score 34; DB 10; Length 589;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 661 ggttcgcatgttgaccaggctggtctcgaaac 694  
|||||  
Db 355 GGGTTCCGATGTTGACGAGCTGCTCGAAC 388

RESULT 24  
BE296260/c  
LOCUS  
DEFINITION 601174838F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:3530265 5',

mRNA sequence.  
BE296260 279 bp mRNA EST 28-AUG-1998  
BE296260.1 GI:9179819  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 734)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
CDNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LLCW202 row: 1 column: 10  
High quality sequence stop: 641.  
Location/Qualifiers  
1..734  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3530265"  
/clone\_lib="NIH\_MGC\_17"  
/tissue\_type="rhabdomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: muscle; Vector: pOT87; Site:1: EcoRI;  
Site:2: XhoI; CDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

183 a 175 c 204 g 172 t  
BASE COUNT  
ORIGIN

Query Match 3.3%; Score 34; DB 10; Length 734;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 663 gtttcgcatgttgaccaggctggtctcgaaactc 696  
|||||  
Db 561 GTTTCGCATGTTGACGAGCTGCTCGAACTC 528

RESULT 25  
AI051456/c  
LOCUS  
DEFINITION AI051456 279 bp mRNA EST 28-AUG-1998  
IMAGE:1648223 3' similar to contains Alu repetitive element,, mRNA  
sequence.  
AI051456  
VERSION AI051456.1 GI:3306990  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 279)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 350 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham.

# FEATURES

source  
 1. 279  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1648223"  
 /clone\_lib="Soares\_parathyroid\_tumor\_NbHPA"  
 /tissue\_type="parathyroid tumor"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Organ: parathyroid gland; Vector: pT73D (Pharmacia)  
 ) with a modified polylinker; Site\_1: Not I; Site\_2: Eco  
 RI; 1st strand cDNA was primed with a Not I - oligo(dT)  
 primer  
 [5'-TGTTACCAATCTGAAGTGGCGCGCGACCAATTTTTTTTTTTTTTTT  
 TTTT-3'], double-stranded cDNA was size selected, ligated  
 to Eco RI adapters (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of a modified pT73  
 vector (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid  
 adenomas was kindly provided by Dr. Stephen Marx, National  
 Institute of Diabetes and Digestive and Kidney Diseases,  
 NIH."

BASE COUNT 49 a 68 c 78 g 84 t  
 ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-06;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttcccatgttgaccagctggttcgaactc 696  
 |||||||  
 Db 279 TTTCCCATGTGACCGAGCTGTCGAAC TC 247

# RESULT 26

AW339622 298 bp mRNA EST 31-JAN-2000  
 LOCUS hel5c05.x1 NCI\_CGAP\_CML1 Homo sapiens cDNA clone IMAGE:2919080 3'  
 DEFINITION similar to contains Alu repetitive element; contains element LTR8  
 repetitive element ;, mRNA sequence.  
 ACCESSION AW339622  
 VERSION AW339622.1 GI:6836248  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 298)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Elisabeth Paletta, Jonathan D. Licht, M.D.,  
 Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life  
 Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The  
 I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
 Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Seq primer: -40UP from gibco  
 High quality sequence stop: 287.

# FEATURES

source  
 1. 298  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2919080"  
 /clone\_lib="NCI\_CGAP\_CML1"  
 /tissue\_type="myeloid cells, 18 pooled CML cases, BCR/ABL  
 rearrangement positive, includes both chronic phase and  
 myeloid blast crisis"  
 /lab\_host="DH10B"  
 /note="Organ: whole blood; Vector: pCMV-SPORT6; Site\_1:  
 SalI; Site\_2: NotI; Cloned unidirectionally. Primer:  
 Oligo dT. Library constructed by Life Technologies."

BASE COUNT 45 a 94 c 82 g 77 t  
 ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 8e-06;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggtttcccatgttgaccagg 682  
 |||||||  
 Db 178 TAGTAGAGACGGGTTTCGCATGTTGACCAGG 210

# RESULT 27

AI244758 348 bp mRNA EST 28-JAN-1999  
 LOCUS qj97g12.x1 NCI\_CGAP\_Kid3 Homo sapiens cDNA clone IMAGE:1867462 3'  
 DEFINITION similar to contains Alu repetitive element; , mRNA sequence.  
 ACCESSION AI244758  
 VERSION AI244758.1 GI:3840155  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 348)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Sequencing by: Greg Lennon, Ph.D.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 304 Std Error: 0.00  
 Seq primer: -40UP from Gibco.

# FEATURES

source  
 1. 348  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1867462"  
 /clone\_lib="NCI\_CGAP\_Kid3"  
 /lab\_host="DH10B"  
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with  
 a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer,  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pT73 vector. mRNA  
 source: 2 pooled kidneys. Library went through one round  
 of normalization. Library constructed by Bento Soares and  
 M. Fatima Bonaldo."

BASE COUNT 70 a 105 c 88 g 85 t  
 ORIGIN



**TITLE** Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

**JOURNAL** Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

**MEDLINE** 99380589

**COMMENT** Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@ejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering.bac.htm>) or from Resear h Genetics (<http://info@resgen.com>). BAC end Web Server: <http://www.htsc.washington.edu>  
Plate: 733 row: N column: 20  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 409.

# FEATURES

**source**

**Location/Qualifiers**

1. .409

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="RPCI-11 Human Male BAC Library"

/sex="male"

/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"

**BASE COUNT** 118 a 105 c 96 g 86 t 4 others

**ORIGIN**

Query Match 3.2%; Score 33; DB 13; Length 409;  
Best Local Similarity 100.0%; Pred. No. 8e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Qy** 664 ttctgccatgttgaccaggctgctcgaaactc 696  
|||||

**Db** 277 ttctgccatgttgaccaggctgctcgaaactc 245  
|||||

**RESULT** 31

**M78087** 411 bp mRNA EST 26-MAY-1992

**LOCUS** EST01677 Subtracted Hippocampus, Stratagene (cat. #936205) Homo sapiens cDNA clone HHCPE75 similar to Alu repetitive element, mRNA sequence.

**DEFINITION** M78087

**ACCESSION** M78087

**VERSION** M78087

**KEYWORDS** EST.

**SOURCE** human.

**ORGANISM** Homo sapiens

**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**AUTHORS** 1 (bases 1 to 411)  
Adams,M.D., Dubnick,M., Kerlavage,A.R., Moreno,R., Kelley,J.M., Utterback,T.R., Nagle,J.W., Fields,C. and Venter,J.C.

**TITLE** Sequence identification of 2,375 human brain genes

**JOURNAL** Nature 355, 632-634 (1992)

**MEDLINE** 92168112

**COMMENT** Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@igr.org  
Seq primer: M13 Forward.

# FEATURES

**source**

**Location/Qualifiers**

1. .411

/organism="Homo sapiens"

/db\_xref="ATCC (inhost):78447"

/db\_xref="GDB:D081086"

/db\_xref="taxon:9606"

/clone\_lib="HHCPE75"

/clone\_lib="Subtracted Hippocampus, Stratagene (cat. #936205)"

/note="Vector: lambdaZAP-II; The hippocampus library (#4 above) was subtracted with a fibroblast cell line cDNA library (Stratagene cat. #936209; WI38 lung fibroblast cell line: oligo-dT + random primed cDNA synthesis; lambdaZAP-II vector, 1.0kb average insert size.) by the method of Sive & St. John (Nucl. Acids Res. 16:10937, 1988 )."

**BASE COUNT** 99 a 114 c 93 g 105 t

**ORIGIN**

Query Match 3.2%; Score 33; DB 11; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Qy** 669 ccattgtgaccaggctgctcgaaactcttgac 701  
|||||

**Db** 63 CCATGTCACCAAGCTGGTCTCGAACTCTTGAC 95  
|||||

**RESULT** 32

**AA457230/c** 417 bp mRNA EST 06-JUN-1997

**LOCUS** AA457230

**DEFINITION** IMAGE:838658 3' similar to contains Alu repetitive element., mRNA sequence.

**ACCESSION** AA457230

**VERSION** AA457230

**KEYWORDS** EST.

**SOURCE** human.

**ORGANISM** Homo sapiens

**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**AUTHORS** 1 (bases 1 to 417)  
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,F., Wylie,T., Waterston,R. and Wilson,R.

**TITLE** Washu-Merck EST Project 1997

**JOURNAL** Unpublished (1997)

**COMMENT** Contact: Wilton RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium ([infoimage.llnl.gov](http://infoimage.llnl.gov)) for further information.  
Possible reversed clone: polyt not found  
Seq primer: -41ml3 fwd. ET from Amersham  
High quality sequence stop: 394.

# FEATURES

**source**

**Location/Qualifiers**

1. .417

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="Stratagene fetal retina 937202"

/sex="mixed"

/lab\_host="SOLR (kanamycin resistant)"

/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"



Db 110 CCATGTTGACAGGCTGCTCGAACTCTTGAC 78

RESULT 34  
BF802026/c 450 bp mRNA EST 12-JAN-2001  
LOCUS CM1-C10092-301000-503-h06 C10092 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BF802026  
ACCESSION BF802026  
VERSION BF802026.1 GI:12131015  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 450)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM1&t2=CM1-C10092-  
301000-503-h06&t3=2000-10-30&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 49  
High quality sequence stop: 450.  
Location/Qualifiers  
1. .450  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="C10092"  
/dev\_stage="Adult"  
/note="Organ: colon\_ins; Vector: puc18; Site\_1: Sma1;  
Site\_2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 115 a 109 c 131 g 95 t  
ORIGIN

Query Match 3.2%; Score 33; DB 11; Length 450;  
Best Local Similarity 100.0%; Pred. No. 7.9e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccattgtgaccagctggtctcgaaactcttgac 701  
|||||

Db 330 CCATGTTGACAGGCTGCTCGAACTCTTGAC 298  
|||||

RESULT 35  
AA984191/c 453 bp mRNA EST 27-MAY-1998  
LOCUS am82b06.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone  
DEFINITION IMAGE:1629587 3' similar to contains Alu repetitive element;; mRNA  
sequence.  
ACCESSION AA984191

BASE COUNT 129 a 95 c 83 g 110 t

Query Match 3.2%; Score 33; DB 10; Length 417;  
Best Local Similarity 100.0%; Pred. No. 7.9e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccattgtgaccagctggtctcgaaactcttgac 701  
|||||

Db 347 CCATGTTGACAGGCTGCTCGAACTCTTGAC 315

RESULT 33  
BE146457/c 447 bp mRNA EST 21-JUN-2000  
LOCUS QV0-HT0215-041199-046-b01 HT0215 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BE146457  
ACCESSION BE146457  
VERSION BE146457.1 GI:8609181  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 447)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-QV0-HT0215-041  
199-046-b01&t3=1999-11-04&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 10  
High quality sequence stop: 447.  
Location/Qualifiers  
1. .447  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="HT0215"  
/dev\_stage="Adult"  
/note="Organ: head\_neck; Vector: puc18; Site\_1: Sma1;  
Site\_2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 140 a 93 c 132 g 82 t  
ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 447;  
Best Local Similarity 100.0%; Pred. No. 7.9e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccattgtgaccagctggtctcgaaactcttgac 701  
|||||

```

VERSION AA984191.1 GI:3162716
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 453)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krisman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
White, Y., Wylie, J., Waterston, R., and Wilson, R.
Wasnu-NCI human EST project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: polyt not found
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 452.
FEATURES
source
1. 453
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1629587"
/sex="male"
/tissue_type="schizophrenic brain S-11 frontal lobe"
/dev_stage="34 years old"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Library
constructed from S-11 frontal lobe, male, 34 years old,
50% caucasian, 50% Aleutian. Schizophrenic suicide.
Random primed into EcoRI site of ZAP II Vector. Mass
excised. Custom library. Avg insert length 1.4kb.
Material obtained by Johnston N., Torrey, E.F., Yolken R.,
and the Stanley Neuropathology Consortium - Analysis of
RNAs from the Brains of Individuals with Psychiatric
Diseases (Unpublished) Stanley Neurovirology Laboratory,
Johns Hopkins School of Medicine, Baltimore MD."
BASE COUNT 126 a 100 c 132 g 95 t
ORIGIN

Query Match 3.2%; Score 33; DB 10; Length 453;
Best Local Similarity 100.0%; Pred. No. 7.9e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttgcgcattgtgaccaggctggctcgaaactc 696
|||||
Db 220 TTTGGCAGTTGACCGAGCTGGTCTCGAATC 188

RESULT 36
BF982842/c
LOCUS BF982842 679 bp mRNA EST 23-JAN-2001
DEFINITION 602305602F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4396752 5',
mRNA sequence.
ACCESSION BF982842
VERSION BF982842
KEYWORDS BF982842.1 GI:12385654
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 679)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cga@rs-rcmail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0096 row: d column: 01
High quality sequence stop: 667.
FEATURES
source
1. 679
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4396752"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 208 a 135 c 160 g 176 t
ORIGIN

Query Match 3.2%; Score 33; DB 11; Length 679;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttgcgcattgtgaccaggctggctcgaaactc 696
|||||
Db 489 TTTGGCAGTTGACCGAGCTGGTCTCGAATC 457

RESULT 37
AQ595304/c
LOCUS AQ595304 712 bp DNA GSS 08-JUN-1999
DEFINITION HS_5448_A2_G08_T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=1024 Col=16 Row=M, DNA sequence.
ACCESSION AQ595304
VERSION AQ595304.1 GI:5026890
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 712)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1024 row: M column: 16
Seq primer: T7
Class: BAC ends

```

us-09-671-050-9.oli20.rst

Fri May 3 11:02:12 2002

High quality sequence stop: 712.  
 Location/Qualifiers  
 1..712  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="Plate=1024 Col=16 Row=M"  
 /clone\_lib="RPCI-11 Human Male BAC Library"  
 /sex="male"  
 /note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 Male blood DNA was isolated from one randomly chosen donor  
 and partially digested with a combination of EcoRI and  
 EcoRI Methylase. Size selected DNA was cloned into the  
 pBACE3.6 vector at EcoRI sites"  
 223 a 141 c 157 g 166 t 25 others

BASE COUNT  
 ORIGIN

Query Match 3.2%; Score 33; DB 13; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-06;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 coattgtgaccaggctgtctcgaaactcttgac 701  
 |||||  
 Db 354 CCATGTTGACCGCTGGTCTCGAACTCTTGAC 322

RESULT 38  
 BG575508/c 859 bp mRNA EST 10-APR-2001  
 LOCUS 602598919f1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:4707460 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG575508  
 VERSION BG575508.1 GI:13583161  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 859)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: DCTD/DTP  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10577 row: f column: 05  
 High quality sequence stop: 812.

FEATURES  
 source  
 1..859  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4707460"  
 /clone\_lib="NIH\_MGC\_87"  
 /tissue\_type="mammary adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organelle: breast; Vector: pCMV-Sport6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.383 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: This is a NIH\_MGC Library."  
 260 a 158 c 197 g 243 t

BASE COUNT  
 ORIGIN

Query Match 3.2%; Score 33; DB 11; Length 859;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-06;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 664 ttctgcacatgttgaccaggctgtctcgaaactc 696  
 |||||  
 Db 476 TTTCCGCATGTTGACCGCTGGTCTCGAACTC 444

RESULT 39  
 BG912455/c 874 bp mRNA EST 05-JUN-2001  
 LOCUS 602806793f1 NCI\_CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4939102  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BG912455  
 VERSION BG912455.1 GI:14292931  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 874)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: David N. Louis, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10875 row: i column: 23  
 High quality sequence stop: 645.

FEATURES  
 source  
 1..874  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4939102"  
 /clone\_lib="NCI\_CGAP\_Brn67"  
 /tissue\_type="anaplastic oligodendroglioma with lp/19q  
 loss"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organelle: brain; Vector: pCMV-Sport6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. primer: oligo dT.  
 Average insert size 2.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 251 a 182 c 204 g 237 t

BASE COUNT  
 ORIGIN

Query Match 3.2%; Score 33; DB 11; Length 874;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-06;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 664 ttctgcacatgttgaccaggctgtctcgaaactc 696  
 |||||  
 Db 293 TTTCCGCATGTTGACCGCTGGTCTCGAACTC 261

RESULT 40  
 BG151645 235 bp mRNA EST 05-FEB-2001  
 LOCUS nai68c04.xl NCI\_CGAP\_HNI4 Homo sapiens cDNA clone IMAGE:4318471  
 DEFINITION similar to contains Alu repetitive element; contains element MER13  
 MER35 repetitive element ;, mRNA sequence.  
 ACCESSION BG151645  
 VERSION BG151645.1 GI:12663675  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 235)

**AUTHORS**  
**TITLE** NCI/NIDR-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**JOURNAL** National Cancer Institute / National Institute of Dental Research,  
**COMMENT** Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-i@mail.nih.gov](mailto:cgaps-i@mail.nih.gov)  
 Tissue Procurement: John F. Ensley, M.D., Chidehanok Leethanakul,  
 D.D.S.

cDNA Library Preparation: Krizman and Leethanakul Laboratories  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL, send email to:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 211.

**FEATURES**  
 source  
 1. .235  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4318471"  
 /clone\_lib="NCI\_CGAP\_HN14"  
 /tissue\_type="hyperplasia of squamous epithelium"  
 /lab\_host="DH10B"  
 /note="Organ: tongue; Vector: pAMP10; mRNA made from  
 tongue epithelium, cDNA made by oligo-dT priming.  
 Non-directionally cloned into UDG sites. Size-selected on  
 agarose gel, average insert size 500 bp. Primary library.  
 cDNA Library Preparation: David B. Krizman, Ph.D.  
 REFERENCE: Krizman et al. (1996) Cancer Research  
 56:5380-5383."

**BASE COUNT** 44 a 57 c 63 g 71 t

**Query Match** 3.1%; Score 32; DB 11; Length 235;  
**Best Local Similarity** 100.0%; Pred. No. 2.6e-05;  
**Matches** 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 681 ggcgtgctgcgaactcttgacgtcaagtgtac 712  
 ||||||||||||||||||||||||||||||||  
**Db** 198 GCGTGGCTCGAAGCTCTTGACGTCAGTGATC 229

**RESULT** 41  
**AG095034**  
**LOCUS** AG095034 392 bp DNA GSS 27-AUG-1998  
**DEFINITION** HS 3028\_A2\_G11\_WF CIT Approved Human Genomic Sperm Library D Homo  
 sapiens genomic clone Plate=3028 Col=22 Row=M, DNA sequence.  
**ACCESSION** AG095034  
**VERSION** AG095034.1 GI:3466451  
**KEYWORDS** GSS.  
**SOURCE** human.

**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 392)  
**AUTHORS** Kahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
 Ketter, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and  
 Hood, L.

**TITLE** Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
**JOURNAL** Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
**MEDLINE** 99380589  
**COMMENT** Contact: Mahairas GG, Wallace JC, Hood L  
 High throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: [jwallace@u.washington.edu](mailto:jwallace@u.washington.edu)  
 Sequence Tagged Connector  
 Plate: 3028 row: M column: 22

**Class:** BAC ends  
**High quality sequence stop:** 392.  
**Location/Qualifiers**  
 1. .392  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="Plate=3028 Col=22 Row=M"  
 /clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /sex="male"  
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in  
 E-Coli DH10B"

**BASE COUNT** 85 a 80 c 85 g 141 t 1 others  
**ORIGIN**

**Query Match** 3.1%; Score 32; DB 13; Length 392;  
**Best Local Similarity** 100.0%; Pred. No. 2.6e-05;  
**Matches** 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 668 gccatgttcacaggctgctgcgaactcttg 699  
 ||||||||||||||||||||||||||||||||  
**Db** 361 GCCATGTTGACCAAGGCTGCTCGAAGCTCTTG 392

**RESULT** 42  
**AG002404/c**  
**LOCUS** AG002404 616 bp DNA GSS 06-FEB-1999  
**DEFINITION** Homo sapiens genomic DNA, 21q region, clone: 149C3X10, genomic  
 survey sequence.  
**ACCESSION** AG002404  
**VERSION** AG002404.1 GI:2627056  
**KEYWORDS** GSS.  
**SOURCE** Homo sapiens DNA, clone:149C3X10.  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**REFERENCE** 1 (bases 1 to 616)  
**AUTHORS** Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (18-NOV-1997) to the DDBJ/EMBL/GenBank databases.  
 Masahira Hattori, Kitasato University, Department of Science, JST  
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
 (E-mail: [hattori@hgc.ims.u-tokyo.ac.jp](mailto:hattori@hgc.ims.u-tokyo.ac.jp), Tel:0427-78-9732,  
 Fax:0427-78-9561)  
**REFERENCE** 2 (bases 1 to 616)  
**AUTHORS** Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
**TITLE** Homo sapiens genomic DNA, chromosome 21q  
**JOURNAL** Published Only in Database (1997) In press  
**FEATURES** Location/Qualifiers  
 source  
 1. .616  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /clone="149C3X10"  
 /map="21q"

**BASE COUNT** 171 a 125 c 161 g 155 t 4 others  
**ORIGIN**

**Query Match** 3.1%; Score 32; DB 13; Length 616;  
**Best Local Similarity** 100.0%; Pred. No. 2.6e-05;  
**Matches** 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 658 acgggggttcgacatgttgaccaggctgtct 689  
 ||||||||||||||||||||||||||||||||  
**Db** 140 ACGGGTTTCCCATGTTGACCAAGCTGGTCT 109

**RESULT** 43  
**AG002403/c**  
**LOCUS** AG002403 658 bp DNA GSS 06-FEB-1999  
**DEFINITION** Homo sapiens genomic DNA, 21q region, clone: 149C3X10, genomic  
 survey sequence.

[illegible]

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 691)  
Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.  
Direct Submission  
Submitted (18-NOV-1997) to the DDBJ/EMBL/GenBank databases.  
Masahira Hattori, Kitasato University, Department of Science, JST  
Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228, Japan  
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,  
Fax:0427-78-9561)  
2 (bases 1 to 691)  
Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.  
Homo sapiens genomic DNA, chromosome 21q  
Published Only in Database (1997) In press  
Location/Qualifiers  
source  
1. .691  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/clone="149C3X10"  
/map="21q"

BASE COUNT 204 a 163 c 124 g 200 t  
ORIGIN

Query Match 3.1%; Score 32; DB 13; Length 691;  
Best Local Similarity 100.0%; Pred. No. 2.6e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 658 acggggtttcgccatgttgaccagctggtct 689  
|||||  
Db 343 ACGGGTTTCGCCATGTTGACCAGCTGCTCT 374  
|||||

RESULT 47  
AA0742378  
LOCUS  
DEFINITION  
HS-382-BL-C09-T7A RPCI-11 Human Male BAC Library Homo sapiens  
genomic clone Plate=958 Col=17 Row=F, DNA sequence.  
AA0742378  
VERSION  
A0742378.1 GI:5519900  
KEYWORDS  
GSS.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 853)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
99380589  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm)  
or from Research h Genetics (info@resgen.com). BAC end Web Server:  
http://www.htsc.washington.edu  
plate: 958 row: F column: 17  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 853.  
Location/Qualifiers  
source  
1. .853  
/organism="Homo sapiens"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 853)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
99380589  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm)  
or from Research h Genetics (info@resgen.com). BAC end Web Server:  
http://www.htsc.washington.edu  
plate: 958 row: F column: 17  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 853.  
Location/Qualifiers  
source  
1. .853  
/organism="Homo sapiens"

/db\_xref="taxon:9606"  
/clone="Plate=958 Col=17 Row=F"  
/clone\_lib="RPCI-11 Human Male BAC Library"  
/sex="male"  
/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBACE3.6 vector at EcoRI sites"  
BASE COUNT 182 a 220 c 181 g 270 t  
ORIGIN

Query Match 3.1%; Score 32; DB 13; Length 853;  
Best Local Similarity 100.0%; Pred. No. 2.5e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 670 catgttaccagctggtctcgaaactcttgaac 701  
|||||  
Db 477 CATGTTGACCAGGCTGGTCTGAACTCTTGAC 508  
|||||

RESULT 48  
AA640211  
LOCUS  
DEFINITION  
AA640211 248 bp mRNA EST 23-OCT-1997  
nn95h08.s1 NCI-CGAP-Phel Homo sapiens cDNA clone IMAGE:1098975 3'  
similar to contains Alu repetitive element; , mRNA sequence.  
AA640211  
ACCESSION  
AA640211.1 GI:2563990  
VERSION  
EST.  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 248)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-r@mail.nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Stratagene, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.lini.gov/bbrp/image/image.html  
Seq primer: -40m13 fwd. Et from Amersham.  
Location/Qualifiers  
source  
1. .248  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1098975"  
/clone\_lib="NCI-CGAP-Phel"  
/tissue\_type="pheochromocytoma"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Vector: Bluescript SK-; Site\_1: EcoRI; Site\_2: XhoI  
; Cloned unidirectionally. Primer: Oligo dT  
Pheochromocytoma. 5' adaptor sequence: 5' GAATTCGCGCAGCAG  
3' 3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTT 3'.  
Average insert size: 1.3 kb."  
BASE COUNT 46 a 69 c 63 g 70 t  
ORIGIN

FEATURES  
source

Query Match 3.0%; Score 31; DB 10; Length 248;  
Best Local Similarity 100.0%; Pred. No. 8.6e-05;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 660 ggggttcgccatgttgaccagctggtctc 690  
|||||

REFERENCE 1 (bases 1 to 439)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-r@mail.nih.gov  
Life Technologies catalog #: 11548-013  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/bbrp/image/image.html  
Insert Length: 674 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 415.  
Location/Qualifiers  
FEATURES  
source  
1..439  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2420367"  
/clone\_lib="NCI-CGAP\_Pan1"  
/tissue\_type="adenocarcinoma"  
/lab\_host="DH10B"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.72 kb. Life Technologies catalog #:  
11548-013"  
BASE COUNT 95 a 115 c 92 g 136 t 1 others  
ORIGIN  
Query Match 3.0%; Score 31; DB 10; Length 439;  
Best Local Similarity 100.0%; Pred. No. 8.4e-05;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 660 ggggttcgccatgttgaccagcgtgtctc 690  
|||||  
Db 197 ggggttcgccatgttgaccagcgtgtctc 227  
Search completed: May 2, 2002, 19:12:32  
Job time: 7541 sec

Db 167 GGGTTTCGCCATGTTGACCAGGCTGTCTC 197  
RESULT 49  
LOCUS AQ341086/c 349 bp DNA GSS 06-MAY-1999  
DEFINITION AQ341086.1 TV RPCI-11 Homo sapiens genomic clone RPCI-11-119H23,  
DNA sequence.  
ACCESSION AQ341086  
VERSION AQ341086.1 GI:4165982  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 349)  
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter  
J.C.  
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building (1997)  
JOURNAL Unpublished (1997)  
COMMENT Other GSSs: RPCI11-119H23.TJ  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/tldb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: T7  
Class: BAC ends.  
Location/Qualifiers  
FEATURES  
source  
1..349  
/organism="Homo sapiens"  
/db\_xref="GDB:7545502"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-119H23"  
/clone\_lib="RPCI-11"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"  
BASE COUNT 112 a 74 c 72 g 91 t  
ORIGIN  
Query Match 3.0%; Score 31; DB 13; Length 349;  
Best Local Similarity 100.0%; Pred. No. 8.5e-05;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 651 agtagagacgggttcgccatgttgaccag 681  
|||||  
Db 284 AGTAGAGACGGGTTTCGCCATGTTGACCAG 254  
RESULT 50  
LOCUS AI814697 439 bp mRNA EST 07-MAR-2000  
DEFINITION AI814697.1 NCI-CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2420367 3',  
similar to contains Alu repetitive element,, mRNA sequence.  
ACCESSION AI814697  
VERSION AI814697.1 GI:5425912  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

.....

---







12	784.5	47.2	591	22	AAU03524	Human protein kina
13	763.5	45.9	457	21	AAU03524	Rat lost in leukae
14	763.5	45.9	505	21	AAU03524	Rat lost in leukae
15	656	39.5	534	22	AAU03524	Novel protein kina
16	581	35.0	333	22	AAU03524	Human polypeptide
17	574	34.5	352	22	AAU03524	Human polypeptide
18	570	34.3	298	20	AAU03524	Human polypeptide
19	570	34.3	544	19	AAU03524	CDK2-green fluores
20	570	34.3	544	19	AAU03524	Green fluorescent
21	568.5	34.2	294	20	AAU03524	Oryza sativa Cdc2
22	566.5	34.1	294	20	AAU03524	Zea mays protein f
23	565.5	34.0	294	21	AAU03524	Zea mays protein f
24	561.5	33.8	294	21	AAU03524	Arabidopsis thalia
25	561.5	33.8	294	21	AAU03524	Arabidopsis thalia
26	561.5	33.8	350	21	AAU03524	Arabidopsis thalia
27	557	33.5	297	20	AAU03524	Rattus norvegicus
28	556	33.5	297	20	AAU03524	Homo sapiens Cdc2
29	555.5	33.4	300	20	AAU03524	Pneumocystis carin
30	555	33.4	270	20	AAU03524	Human protein kina
31	554	33.3	298	21	AAU03524	Cyclin-dependent k
32	544	32.7	292	15	AAU03524	Cyclin dependent p
33	542	32.6	292	21	AAU03524	Protein involved i
34	539.5	32.5	294	22	AAU03524	Tomato cyclin-depe
35	518.5	31.2	297	20	AAU03524	Schizosaccharomyce
36	511	30.7	274	20	AAU03524	Fission yeast prot
37	510.5	30.7	311	21	AAU03524	Arabidopsis thalia
38	507	30.5	348	21	AAU03524	Arabidopsis thalia
39	507	30.5	357	21	AAU03524	Arabidopsis thalia
40	505.5	30.4	317	18	AAU03524	Arabidopsis thalia
41	505.5	30.4	317	20	AAU03524	Arabidopsis thalia
42	505.5	30.4	317	21	AAU03524	Arabidopsis thalia
43	505	30.4	323	20	AAU03524	Candida albicans C
44	500	30.1	346	21	AAU03524	Emeritella nidulan
45	498	30.0	306	20	AAU03524	Human cyclin depen
						Soybean cyclin-dep

ALIGNMENTS

RESULT	1	
AAE00495		
ID	AAE00495 standard; Protein; 315 AA.	
XX		
AC	AAE00495;	
XX		
DT	19-JUN-2001 (first entry)	
DE	Human kinase #6.	
XX		
KW	Human; kinase; gene therapy; bioreactor; mental disorder;	
KW	biological disorder.	
XX		
OS	Homo sapiens.	
XX		
PN	WC200123579-Al.	
XX		
PD	05-APR-2001.	
XX		
PF	27-SEP-2000; 2000WO-US26621.	
XX		
PR	28-SEP-1999; 99US-0156511.	
XX		
PA	(LEXI-) LEXICON GENETICS INC.	
XX		
PI	Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;	
XX		
DR	WPI; 2001-266166/27.	
XX		
PT	New isolated human kinase polynucleotide useful for generating for	
PT	antibodies, as reagents in diagnostic assays and for screening for	
PT	compounds useful for treating mental, biological or medical diseases	
XX		

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:02:15 ; Search time 40.16 Seconds  
(without alignments)  
581.003 Million cell updates/sec

Title: US-09-671-050-12

Perfect score: 1662

Sequence: 1 MEKYEKLAKTGEGSYGVFK.....RKARNEGRNRQVLPKLS 315

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1662	100.0	315	22	AAE00495
2	1638	98.6	324	22	AAE00492
3	1636	98.4	347	22	AAE00494
4	1619	97.4	360	22	AAU03525
5	1612	97.0	356	22	AAU03525
6	1296.5	78.0	296	22	AAU03525
7	1149	69.1	247	22	AAU03525
8	967.5	58.2	566	21	AAU0724
9	791	47.6	187	22	AAE00490
10	791	47.6	198	22	AAE00493
11	784.5	47.2	455	21	AAU07126





CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
CC Additionally, polynucleotides encoding protein kinases may be  
CC used for gene therapy and as DNA probes in diagnostic assays.  
CC The protein kinase polypeptides may be used as antigens in the production  
CC of antibodies against the protein kinases and in assays to identify  
CC modulators of protein kinase expression and activity.  
XX  
SQ Sequence 360 AA;

Query Match 97.4%; Score 1619; DB 22; Length 360;  
Best Local Similarity 93.4%; Pred. No. 1.8e-164;  
Matches 311; Conservative 2; Mismatches 0; Indels 20; Gaps 2;  
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGOVAVKFFVESEDDPVVKKIATREIRMLKQLKH 60  
DB 1 mekyeklaktgeisgyvfvkcrnktsggvavkffveseddpvkkialreirmlkqlkh 60  
QY 61 PNLVNLIEVFRKRKMHILVFECYDHTLLNELRPNPGVADGVKISVLWQTLQALNFCCHI 120  
DB 61 pnlvnlievfrkrkmhylvfecydhltllnelernpngvadgvkislwqtlqalnfcchih 120  
QY 121 NCIHRRDKPENILITKQGIKICDFGFAQIL-----IPGDAYTDYVA 162  
DB 121 ncihrdkpenilitkqgiikicdfgfaqil :|||||  
QY 163 TRWYRAPELLVGDYQSSYVDIWAIGCVFAELLTGQPLWPGKSDVDQIYLIIRTLGKLIP 222  
DB 181 trwyrapellvgdyqssyvdwaigcvfaelltgqplwpgksdvdqilyliirtlgklip 240  
QY 223 RHQSIFKNGFFHGISIPEPEDMETLEEKFSVHPVALNFMKGCLKYNPDRLTCSQILLE 282  
DB 241 rhqsifkngffhgisipepedmetleekfsdvhvvalnfmkgclkmnpdrltcsqille 300  
QY 283 SSYFDSFOEAQIKRKARNRNRQ--QVLPL 313  
DB 301 ssyfsdsfqaqikrkarnegrrrrrqgnqlpl 333

RESULT 5  
AAE00491  
ID AAE00491 standard; Protein; 356 AA.  
XX  
AC AAE00491;  
XX  
DT 19-JUN-2001 (first entry)  
XX  
DE Human kinase #2.  
XX  
KW Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200123579-A1.  
XX  
PD 05-APR-2001.  
XX  
PF 27-SEP-2000; 2000WO-US26621.  
XX  
PR 28-SEP-1999; 99US-0156511.  
XX  
PA (LEXI-) LEXICON GENETICS INC.  
XX  
PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI; 2001-266166/27.  
DR N-PSDB; AAD03813.  
XX

New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases -  
XX

PS Disclosure; Page 28-29; 38pp; English.  
XX  
CC The present sequence is novel human protein  
CC (NHP) known as human kinase. The human kinases share structural  
CC similarity with animal kinases, more particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.  
XX  
SQ Sequence 356 AA;

Query Match 97.0%; Score 1612; DB 22; Length 356;  
Best Local Similarity 90.6%; Pred. No. 1e-163;  
Matches 310; Conservative 0; Mismatches 0; Indels 32; Gaps 1;  
QY 1 MEKYEKLAKTGEISYGVVFKCRNKTSGOVAVKFFVESEDDPVVKKIATREIRMLKQLKH 60  
DB 1 mekyeklaktgeisgyvfvkcrnktsggvavkffveseddpvkkialreirmlkqlkh 60  
QY 61 PNLVNLIEVFRKRKMHILVFECYDHTLLNELRPNPGVADGVKISVLWQTLQALNFCCHI 120  
DB 61 pnlvnlievfrkrkmhylvfecydhltllnelernpngvadgvkislwqtlqalnfcchih 120  
QY 121 NCIHRRDKPENILITKQGIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDYQGS 180  
DB 121 ncihrdkpenilitkqgiikicdfgfaqilipgdaytdyvatrwyrapellvgdyqgs 180  
QY 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQIYLIIRTL----- 217  
DB 181 svidwaigcvfaelltgqplwpgksdvdqilyliirtlvetgfrhvdqaglleltssdppa 240  
QY 218 -----GKLIPIHQSIKSNCFPHGISIPEPEDMETLEEKFSVHPVALNFMKGCLK 268  
DB 241 vasqsagitgkliprhgsifkngffhgisipepedmetleekfsdvhvvalnfmkgclk 300  
QY 269 MNPDDRITCSQILLESSYFDSFOEAQIKRKARNRNRNRQV 310  
DB 301 mnpddrltcsqillessyfsdsfqaqikrkarnegrrrrrqv 342

RESULT 6  
AAB65643  
ID AAB65643 standard; Protein; 296 AA.  
XX  
AC AAB65643;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE Novel protein kinase, SEQ ID NO: 170.  
XX  
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;  
KW dermatological; antidiabetic; antinfertility; gene therapy; vaccine;  
KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;

KW inflammatory pelvic disease; multiple sclerosis; psoriasis.  
XX Mus musculus.  
OS  
XX WQ200073469-A2.  
PN  
XX 07-DEC-2000.  
PD  
XX  
XX 26-MAY-2000; 2000WO-US14842.  
PF  
XX  
XX 28-MAY-1999; 99US-0136503.  
PR  
XX  
XX (SUG-) SUGEN INC.  
PA  
XX  
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;  
PI  
XX WPI; 2001-032161/04.  
XX N-PSDB; AAF44670.  
XX  
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
DR treating immune-related diseases and disorders, cardiovascular disease,  
DR neurodegenerative diseases and/or cancers -  
XX  
XX Claim 10; Fig 1; 310pp; English.  
XX  
XX The present sequence is a novel protein kinase. The novel protein kinases  
CC and the nucleic acids that encode them may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
CC reproductive disorders.  
XX  
SQ Sequence 296 AA;  
  
Query Match 78.0%; Score 1296.5; DB 22; Length 296;  
Best Local Similarity 90.0%; Pred. No. 4.2e-130;  
Matches 243; Conservative 19; Mismatches 5; Indels 3; Gaps 2;  
  
QY 46 KIALRETRMLKQLKHPNLVNLIEVFRKRKMLHVPYCDHTLLNELRPNPGVADGVKS 105  
DB 1 kialretrmlk-1khpnlvnlievfrkrkmlhvfeycdhtllnelrpnpgvsgdvks 59  
QY 106 VLWQTQLALNFCIHNCIHRDIKPNILITKGGIIKICDFGFAQLIPGDYTDVATRW 165  
DB 60 vlwtqlalnfcchkhncihrdvknpenilitkggmikicdfgarilipgdtydvatrw 119  
QY 166 YRAPPELLVGTQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLKLIIPRHQ 225  
DB 120 yrapellvgdtygssvdvawgcvfaelltgqplwpgksavdqllylirtlkgliiprhq 179  
QY 226 SIFKSNFGFFHGISPEPEDMETLEEKFSVHPVALNFWKGLCKMNPDDRILTCQSLESY 285  
DB 180 sifrsnfgfsgispepedmetleekfsnvqpvalsfmgkglckmnpdrilctcaqildsay 239  
QY 286 FDSFOEAQIKRKARNRGRNRQ--QVLPL 313  
DB 240 festrqedgmkrkarsgrrrrqnqlpl 269  
  
RESULT 7  
AAB65642

ID AAB65642 standard; Protein; 247 AA.  
XX  
AC AAB65642;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE Novel protein kinase, SEQ ID NO: 169.  
XX  
XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;  
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;  
KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.  
XX  
OS Homo sapiens.  
XX  
XX WQ200073469-A2.  
PN  
XX 07-DEC-2000.  
PD  
XX  
XX 26-MAY-2000; 2000WO-US14842.  
PF  
XX  
XX 28-MAY-1999; 99US-0136503.  
PR  
XX  
XX (SUG-) SUGEN INC.  
PA  
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;  
PI  
XX WPI; 2001-032161/04.  
XX N-PSDB; AAF44669.  
XX  
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers -  
XX  
XX Claim 10; Fig 1; 310pp; English.  
XX  
XX The present sequence is a novel protein kinase. The novel protein kinases  
CC and the nucleic acids that encode them may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC expression and activity. Diseases related to kinase expression and  
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
CC reproductive disorders.  
XX  
SQ Sequence 247 AA;  
  
Query Match 69.1%; Score 1149; DB 22; Length 247;  
Best Local Similarity 98.6%; Pred. No. 1.9e-114;  
Matches 217; Conservative 1; Mismatches 0; Indels 2; Gaps 1;  
  
QY 96 NGVADGVKSVLWQTLQALNFCIHNCIHRDIKPNILITKGGIIKICDFGFAQLIPGD 155  
DB 1 ngvadgvksvlwqtlqalnfcchkhncihrdvknpenilitkgiikicdfgarilipgd 60  
QY 156 AYTDYVATRWYRAPPELLVGTQYSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIR 215  
DB 61 aytdyvatrwyrappellvgdtygssvdiwaigcvfaelltgqplwpgksdvdllylir 120  
QY 216 TLGKLIIPRHQSIFKSNFGFFHGISPEPEDMETLEEKFSVHPVALNFMKGLCKMNPDDRIL 275  
DB 161 tlgkliiprhqsifksgfghgispepedmetleekfsdvhpvalnfmkglckmnpddrl 111

Db 121 tlgkliprhqsfkngffhgisipepedmetleekfsdvhvvalnfmkgclkmppddl 180  
 QY 276 TCSOLLESYFDSFOEQAQIKRKARNRNRQ--QVLPL 313  
 Db 181 tcsqlesyfsfdeaqikrkarnegrnrrrqnuqlpl 220

## RESULT 8

AAE00724

ID AAY90724 standard; Protein; 566 AA.

AC AAY90724;

XX

DT 15-AUG-2000 (first entry)

DE Rabbit KKIAMRE kinase SEQ ID NO:4.

XX Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;

KW cdc2-related kinase; Brain; gene therapy; genetic disorder; detection;

KW identification.

XX Oryctolagus cuniculus.

OS WO200020567-A2.

PN 13-APR-2000.

XX 01-OCT-1999; 99WO-US23010.

PP 02-OCT-1998; 98US-0102906.

PR (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX Thompson RF, Gomi H, Sun W;

XX WPI: 2000-328932/28.

XX N-PSDB; AAA29745.

XX Novel learning induced kinase polynucleotides and polypeptides, useful

PT for the analysis of learning and memory, and for gene therapy -

XX Claim 1; Fig 4; 64pp; English.

XX The present sequence represents a learning-induced kinase, designated

CC KKIAMRE kinase which is isolated from rabbit brain tissue. KKIAMRE

CC kinase is a cdc2-related kinase. The KKIAMRE kinase polynucleotides can

CC be used to express recombinant protein for analysis, characterisation or

CC therapeutic use, as markers for tissues in which the protein is

CC preferentially expressed, as molecular weight markers on Southern gels,

CC as chromosome markers or tags, to compare endogenous DNA sequences in

CC patients to identify potential genetic disorders, as probes to hybridise

CC and discover novel related sequences, as a source of PCR primers, and as

CC an antigen to induce anti-DNA antibodies. The polypeptides can be used

CC in assay to discover biological activity, to raise antibodies, as tissue

CC markers, and to isolate correlative receptors or ligands. The

CC polynucleotides may also be used for gene therapy for the treatment of

CC disorders which are mediated by KKIAMRE kinase.

XX Sequence 566 AA;

SQ

## Query Match

Best Local Similarity 58.2%; Score 967.5; DB 21; Length 566;

Matches 177; Conservative 67; Mismatches 65; Indels 15; Gaps 5;

QY 1 MEKYELATGGSGVGVVFKCRNKTSGQVAVYKFFVSEDDPVVKIALREIRMLKOLKH 60

Db 1 mekyenlgivgeysgmvmkcrnkdgrivaikflesdddkmvmkiamreiklqlrh 60

QY 61 PNLVNLIEVFRKRMHLVREYCDHTLLNELRPNPVGADVTKSVLWQTLQALNFCIH 120

Db 61 enlvnllevckkkrrwlvfefvdhtildllefngldqdvqvqylfingifchsh 120

QY 121 NCIHRRDKPENILITKQGIKICDFQAQIL-IPGDAYTDYVATRWYRAPPELLVGDQYX 179  
 Db 121 nihrdikenilvsgvkvkicdfgfartlaapgevydyvatrwyrappellvgdvkyg 180  
 QY 180 SSVDIWAICVFAELLTGQPLWPGKSDVDQYLIIRTLGLIPRHQSIFKSNGFHGISI 239  
 Db 181 kavdvwaigclvtemlmgeplfpgdsdidqlylimrcnlprhqelfyknvfvagvrl 240  
 QY 240 PEPEDMETLEEKFSDVHPVALNFMKCLKMPPDRRLTCSQLLESYF--DSF-----QEA 292  
 Db 241 peikeseperrypkisevvidlakklhvdpdkrfpcaellhhdfqmgdgaerfsqel 300  
 QY 293 QIK--RKARN-----EGRNRRRQ 309  
 Db 301 qmkvqkdarnisiskksqrkkek 324

## RESULT 9

AAE00490

ID AAE00490 standard; Protein; 187 AA.

XX AC AAE00490;

XX 19-JUN-2001 (first entry)

XX Human kinase #1.

XX Human; kinase; gene therapy; bioreactor; mental disorder;

XX biological disorder.

XX Homo sapiens.

XX WO200123579-A1.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26621.

XX 28-SEP-1999; 99US-0156511.

XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX WPI: 2001-266166/27.

XX N-PSDB; AAD03812, AAD03818.

XX New isolated human kinase polynucleotide useful for generating

PT antibodies, as reagents in diagnostic assays and for screening for

PT compounds useful for treating mental, biological or medical diseases -

XX Disclosure; Page 27-28; 38pp; English.

XX The present sequence is novel human protein (NHP) known

CC as human kinase. The human kinases share structural

CC similarity with animal kinases, more particularly serine or

CC threonine protein kinases. Human kinase cDNA is useful for the

CC detection of mutant human kinase for the diagnosis of disease,

CC and also as a therapeutic. It is useful for screening drugs

CC effective in the treatment of symptomatic or phenotypic

CC manifestations perturbing the normal function of NHP in the

CC body. The NHP nucleotide sequences are useful for generation of

CC antibodies, as reagents in diagnostic assays, for the

CC identification of other cellular gene products related to human

CC kinases, and as reagents in assays for screening compounds that

CC are useful for treating mental, biological or medical disorders.

CC NHP oligonucleotides are used as probes. The labelled NHP probes

CC are useful for screening human genomic library for identifying

CC polymorphisms and as primers in amplification assays to detect

CC mutations within the exons, introns and splice sites that can

CC be used in diagnostics and pharmacogenomics. Nucleotide construct

CC encoding NHP products are used to genetically engineer cells

CC in vivo that functions as bioreactors in the body delivering a





PI Zanke B, Haq R, Randall S, Midmer M;  
 XX WPI: 2000-237880/20.  
 DR N-PSDB; AAS21208.  
 XX  
 XX Isolated polynucleotide encoding a lost in leukemia kinase (LLK)  
 PT protein, useful for treatment, diagnosis and prevention of leukemia -  
 XX  
 PS Claim 17; Page 51; 69pp; English.  
 XX  
 CC The present amino acid sequence is the human lost in leukaemia kinase  
 CC (LLK), isolated from human lambda gt10 foetal heart genomic library.  
 CC This gene is mapped to chromosome 5q31.1, a region implicated in human  
 CC acute leukaemia. It is expressed strongly in muscle, heart, liver, brain  
 CC and in tissues which have a very low mitotic index. Human LLK has 83%  
 CC amino acid sequence homology to rat LLK beta protein. LLK is closely  
 CC related to cdc2-related kinases, that are putative tumour suppressors  
 CC and to mitogen activated protein kinases (MAPKs). This sequence has  
 CC cytotatic activity. It is useful for the treatment, diagnosis and  
 CC prevention of acute leukaemia and is also used in gene therapy.  
 XX  
 SQ Sequence 455 AA;

Query Match 47.2%; Score 784.5; DB 21; Length 455;  
 Best Local Similarity 48.5%; Pred. No. 4.6e-75;  
 Matches 150; Conservative 61; Mismatches 91; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGOVAVKVFVESEDDPVVKKIALREIRMLKOLKH 60  
 DB 1 memyetlgkvgegygtvmckkhtgtivaikfyerpqgs-vnkiamreikflkqfhh 59  
 QY 61 PNLVNLIEVFRKRMHLVFEYCDHTLLNELRNPNMGVADGVKSVLWOTLQALNFCIH 120  
 DB 60 enlvnlievfrqkklhivfefidhtvldelqhychgeskrirkyflqilraiyhsh 119  
 QY 121 NCIHRIKPENILITKQIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDQYQ 179  
 DB 120 nihrdikpenilvsgsgitkldcfgfartlaapgdlydyvatrwyrapelvldksy 179  
 QY 180 SSVDAWAGCVFAELLTGQPLWPKGSDVDQLYLIIRTLGKLIPRHQSIFKSNFGFHGISI 239  
 DB 180 kpvdiwaigcmliematgmpylpsssdldlhkvlkvgnlsphlqnifskspifagvvl 239  
 QY 240 PEPDMETLEEFSDVHPVAFNFMKGLKMNPDRLTCSOLLESSYF--DSFQ8---AQI 294  
 DB 240 pqvghpknarkkypkinglladivhaciqdpadrissdilhheyfrdgyfiekfmpel 299  
 QY 295 KRKARNEGR 303  
 DB 300 kakllqea 308

RESULT 12  
 AAU03524  
 ID AAU03524 standard; Protein; 591 AA.  
 XX  
 AC AAU03524;  
 XX  
 DT 12-SEP-2001 (first entry)  
 XX  
 DE Human protein kinase #24.  
 XX  
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
 KW metabolic disorder; immune related disease; neurological disorder;  
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
 KW reproductive disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200138503-A2.  
 XX  
 PD 31-MAY-2001.

XX  
 PF 22-NOV-2000; 2000WO-US32085.  
 XX  
 PR 24-NOV-1999; 99US-0167482.  
 XX  
 PA (SUGF-) SUGEN INC.  
 XX  
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
 PI Flanagan P, Clary D;  
 XX  
 XX WPI: 2001-343950/36.  
 DR N-PSDB; AAS06724.  
 DR  
 XX  
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing  
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
 PT neuronal-associated diseases, and microbial infections -  
 XX  
 XX Claim 7; Figure 2; 433pp; English.  
 PS  
 XX  
 CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The  
 CC novel protein kinases have been identified as members of the tyrosine  
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides  
 CC encoding protein kinases and the polypeptides may be used in the  
 CC prevention, diagnosis and treatment of diseases associated with  
 CC inappropriate kinase expression. For example, they may be used to treat  
 CC cancers (especially cancers of haematopoietic origin), cardiovascular  
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
 CC immune related diseases (e.g. rheumatoid arthritis), neurological  
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
 CC Additionally, polynucleotides encoding protein kinases may be  
 CC used for gene therapy and as DNA probes in diagnostic assays.  
 CC The protein kinase polypeptides may be used as antigens in the production  
 CC of antibodies against the protein kinases and in assays to identify  
 CC modulators of protein kinase expression and activity.  
 XX  
 SQ Sequence 591 AA;

Query Match 47.2%; Score 784.5; DB 22; Length 591;  
 Best Local Similarity 48.5%; Pred. No. 6.9e-75;  
 Matches 150; Conservative 61; Mismatches 91; Indels 7; Gaps 4;  
 QY 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGOVAVKVFVESEDDPVVKKIALREIRMLKOLKH 60  
 DB 1 memyetlgkvgegygtvmckkhtgtivaikfyerpqgs-vnkiamreikflkqfhh 59  
 QY 61 PNLVNLIEVFRKRMHLVFEYCDHTLLNELRNPNMGVADGVKSVLWOTLQALNFCIH 120  
 DB 60 enlvnlievfrqkklhivfefidhtvldelqhychgeskrirkyflqilraiyhsh 119  
 QY 121 NCIHRIKPENILITKQIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDQYQ 179  
 DB 120 nihrdikpenilvsgsgitkldcfgfartlaapgdlydyvatrwyrapelvldksy 179  
 QY 180 SSVDAWAGCVFAELLTGQPLWPKGSDVDQLYLIIRTLGKLIPRHQSIFKSNFGFHGISI 239  
 DB 180 kpvdiwaigcmliematgmpylpsssdldlhkvlkvgnlsphlqnifskspifagvvl 239  
 QY 240 PEPDMETLEEFSDVHPVAFNFMKGLKMNPDRLTCSOLLESSYF--DSFQ8---AQI 294  
 DB 240 pqvghpknarkkypkinglladivhaciqdpadrissdilhheyfrdgyfiekfmpel 299  
 QY 295 KRKARNEGR 303  
 DB 300 kakllqea 308

RESULT 13  
 AAU0125  
 ID AAY70125 standard; Protein; 457 AA.  
 XX

AC AAY70125;  
XX  
DT 06-JUN-2000 (first entry)  
XX  
DE Rat lost in leukaemia kinase beta isoform (LLK beta).  
XX  
KW Lost in leukaemia kinase; LLK beta; rat; cdc2-related kinase; leukaemia;  
KW MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;  
KW cytosstatic; mitotic index; treatment; prevention; gene therapy.  
XX  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT Region 14..15  
FT /label= MAP\_kinase-like\_regulatory\_region  
FT /note= "SY duplex"  
FT Binding-site 44..50  
FT /label= Cyclin\_binding\_motif  
FT Region 125..130  
FT /label= Consensus\_sequence  
FT /note= "Serine/threonine specific kinase sequence"  
FT Region 158..160  
FT /label= MAP\_kinase-like\_regulatory\_region  
FT Region 162..167  
FT /label= Consensus\_sequence  
FT /note= "Serine/threonine specific kinase sequence"  
XX WO200012719-A1.  
XX  
XX PD 09-MAR-2000.  
XX  
XX 31-AUG-1999; 99WO-CA00794.  
XX  
XX 31-AUG-1998; 98CA-2243784.  
XX  
XX 20-NOV-1998; 98CA-2251249.  
XX  
XX (ONTA-) ONTARIO CANCER INST.  
XX  
XX Zanke B, Haq R, Randall S, Midmer M;  
XX WPI; 2000-237880/20.  
XX N-PSDB; AAZ51207.  
XX  
XX Isolated polynucleotide encoding a lost in leukemia kinase (LLK)  
XX protein, useful for treatment, diagnosis and prevention of leukemia -  
XX  
XX Claim 15; Page 48; 69pp; English.  
XX  
XX The present amino acid sequence is the beta isoform of rat lost in  
XX leukaemia kinase (LLK beta), isolated from rat jejunum cDNA library. It  
XX is expressed strongly in muscle, heart, liver, brain and in tissues which  
XX have a very low mitotic index. Rat LLK beta protein is closely related to  
XX cdc2-related kinases, that are putative tumour suppressors and to mitogen  
XX activated protein kinases (MAPKs). This sequence has cytosstatic activity.  
XX It is useful for the treatment, diagnosis and prevention of acute  
XX leukaemia and is also used in gene therapy.  
XX  
XX Sequence 457 AA;  
XX  
Query Match 45.9%; Score 763.5; DB 21; Length 457;  
Best Local Similarity 46.9%; Pred. No. 8.2e-73;  
Matches 145; Conservative 65; Mismatches 92; Indels 7; Gaps 4;  
QY 1 MEKYEKIAKTGSGVGVFKCRNKTSGQVAVKFESEDDPVVKKIALRETRMLKOLKH 60  
DB 1 memyetlgkvgegsygtvmckhkdgrivaikfiye-kpeksvknklatrelkfkqfrh 59  
QY 61 PNLVNLIEVFRKRKXHLVFFEYCDHTLLNLELRNPNGVADGVKTSVWQTLQALNFCIH 120  
DB 60 enlvnlievfrqkkihvlvfeidhtvldelqhnycghieskrkylfqlrateylhnn 119  
QY 121 NCIHREDIKPENILITKQGIKICDFGFAQIL-IPGDATDYVATRWYRAPELLVGDTQYG 179

Db 120 nihrdikpenilvsqgikicdfgfartlaapgdvydyvatrwyrapelvldkdttyg 179  
QY 180 SSVDIWAIGCVFAELLTGOPLWPGKSDVDQLYLIIRTLGKLIPRHOSIFKNGFFHGISI 239  
Db 180 kpvdiwalgcmliematngpylpsssdldllhkiivkgvnltpihlfnfskspifagvvl 239  
QY 240 PEPEDMETLEKFSDDVHPVALNFMKCGLKMPDORLTCSSOLLESSYF--DSFQF---AQI 294  
Db 240 pqvghpknarckypkinglladihvaciqpaeicrsistdlldhdyftrdgmfiel 299  
QY 295 KRKARNEGR 303  
Db 300 rakllgeak 308  
RESULT 14  
AAY70124  
ID AAY70124 standard; Protein; 505 AA.  
XX  
AC AAY70124;  
XX  
DT 06-JUN-2000 (first entry)  
XX  
DE Rat lost in leukaemia kinase alpha isoform (LLK alpha).  
XX  
KW Lost in leukaemia kinase; LLK alpha; rat; cdc2-related kinase; leukaemia;  
KW MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;  
KW cytosstatic; mitotic index; treatment; prevention; gene therapy.  
XX  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT Region 14..15  
FT /label= MAP\_kinase-like\_regulatory\_region  
FT /note= "SY duplex"  
FT Binding-site 44..50  
FT /label= Cyclin\_binding\_motif  
FT Region 125..130  
FT /label= Consensus\_sequence  
FT /note= "Serine/threonine specific kinase sequence"  
FT Region 158..160  
FT /label= MAP\_kinase-like\_regulatory\_region  
FT Region 162..167  
FT /label= Consensus\_sequence  
FT /note= "Serine/threonine specific kinase sequence"  
FT Domain 473..477  
FT /label= Nuclear\_localisation\_motif  
FT Region 480..493  
FT /label= PEST\_sequence  
XX WO200012719-A1.  
XX  
XX 09-MAR-2000.  
XX  
XX 31-AUG-1999; 99WO-CA00794.  
XX  
XX 31-AUG-1998; 98CA-2243784.  
XX 20-NOV-1998; 98CA-2251249.  
XX  
XX (ONTA-) ONTARIO CANCER INST.  
XX  
XX Zanke B, Haq R, Randall S, Midmer M;  
XX WPI; 2000-237880/20.  
XX N-PSDB; AAZ51206.  
XX  
XX Isolated polynucleotide encoding a lost in leukemia kinase (LLK)  
XX protein, useful for treatment, diagnosis and prevention of leukemia -  
XX  
XX Claim 15; Page 47; 69pp; English.  
XX  
XX The present amino acid sequence is the alpha isoform of rat lost in

CC leukaemia kinase (LLK alpha), isolated from rat brain cDNA library. It  
CC is expressed strongly in muscle, heart, liver, brain and in tissues which  
CC have a very low mitotic index. Rat LLK alpha protein has 47% and 44%  
CC sequence homology to KIAMRE and KIALRE cdc2-related kinases, that are  
CC putative tumour suppressors and to mitogen-activated protein kinases  
CC (MAPKs). This sequence has cytosstatic activity. It is useful for the  
CC treatment, diagnosis and prevention of acute leukaemia and is also used  
CC in gene therapy.  
XX

SQ Sequence 505 AA;

Query Match 45.9%; Score 763.5; DB 21; Length 505;  
Best Local Similarity 46.9%; Pred. No. 9.5e-73;

Matches 145; Conservative 65; Mismatches 92; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGOVVAVKKEVEDDPVVKIALREIRMLKQLKH 60  
DB 1 MEMYETLGKVGEGSYGVVFKCRNKTSGOVVAVKKEVEDDPVVKIALREIRMLKQLKH 60  
QY 61 PNLVNLIEVFRKRMHILVFYCDHTLLNELERNPNNGVADGVKSVLMOTLQALNFCIH 120  
DB 60 ENLVNLIEVFRKRMHILVFYCDHTLLNELERNPNNGVADGVKSVLMOTLQALNFCIH 120  
QY 121 NCIHDRKIPENILITKOGIKICDFGFAQIL-IPGDVATDVYVATRWYRAPPELLVGDTOY 178  
DB 120 NVIIHRDKIPENILITKOGIKICDFGFAQIL-IPGDVATDVYVATRWYRAPPELLVGDTOY 178  
QY 180 SSVDIWAGCVFAELLTGOPLMWPKGSDVDQLYLIITGLKILPRHQSIFKSGFFHGISI 239  
DB 180 KPVEDMETLEKESDVHPVAFNMFKGLKMFNDRDLTCSQLLESSYF--DSFOE---AQI 294  
QY 240 PEPEDMETLEKESDVHPVAFNMFKGLKMFNDRDLTCSQLLESSYF--DSFOE---AQI 294  
DB 240 PQVGHPNARKKYPKINGLADIVHACLGIDPAERISSTDLIHHGYFTDGFIEKFIPEL 299  
QY 295 KRKARNEGR 303  
DB 300 RAKILQEAQ 308

RESULT 15

AA065645  
ID AA065645 standard; Protein; 534 AA.

AA065645;

27-MAR-2001 (first entry)

Novel protein kinase, SEQ ID NO: 172.

Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;  
dermatological; antidiabetic; antifertility; gene therapy; vaccine;  
immune disorder; cardiovascular disease; neurodegenerative disease;  
cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
inflammatory pelvic disease; multiple sclerosis; psoriasis.

Homo sapiens.

W0200073469-A2.

07-DEC-2000.

26-MAY-2000; 2000WO-US14842.

28-MAY-1999; 99US-0136503.

(SUGEN-) SUGEN INC.

Plowman GD, Martinez R, Whyte D, Sudersanam S;

WPI; 2001-032161/04.

DR N-PSDB; AAF44672.

XX  
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers -  
XX

PS Claim 10; Fig 1; 310pp; English.

XX  
XX The present sequence is a novel protein kinase. The novel protein kinases  
CC and the nucleic acids that encode them may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
XX reproductive disorders.

SQ Sequence 534 AA;

Query Match 39.5%; Score 656; DB 22; Length 534;  
Best Local Similarity 38.0%; Pred. No. 3.3e-61;

Matches 153; Conservative 58; Mismatches 100; Indels 92; Gaps 11;

QY 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGOVVAVKKEVEDDPVVKIALREIRMLKQLKH 60

DB 1 MEMYETLGKVGEGSYGVVFKCRNKTSGOVVAVKKEVEDDPVVKIALREIRMLKQLKH 60

QY 61 PNLVNLIEVFRKRMHILVFYCDHTLLNELERNPNNGVADGVKSVLMOTLQALNFCIH 120

DB 60 ENLVNLIEVFRKRMHILVFYCDHTLLNELERNPNNGVADGVKSVLMOTLQALNFCIH 120

QY 121 NCIHDRKIPENILITKOGIKICDFGFAQIL-IPGDVATDVYVATRWYRAPPELLVGDTOY 178

DB 120 NVIIHRDKIPENILITKOGIKICDFGFAQIL-IPGDVATDVYVATRWYRAPPELLVGDTOY 178

QY 179 GS--SVDIWAIGCVFAELLTGOPLMWPKGSDVDQLYLIITGLKILPRHQSIFKSGFFHGISI 239

DB 180 KPVEDMETLEKESDVHPVAFNMFKGLKMFNDRDLTCSQLLESSYF--DSFOE---AQI 294

QY 215 -----RTGLKILPRHQSIFKSGFFHGISI 239

DB 240 IKPKESSEKENELRKDERKTYTNTLLSSVGLKELEKKEKKEIKVVRVKGGRGDIS 299

QY 240 PEPEDME---TLEKESDVHPVAFNMFKGLKMFNDRDLTCSQLLESSYF--DSFOE---AQI 294

DB 300 PKKKEVEGGLGQDANVHPSMDTKIVTIEPPNPINPSTCNGIKENPHCGGSGVTMP 359

QY 273 DRLTCSQLLESSYFDSFOEQAQIKRKARNEGRNRRO---QVLP 312

DB 360 INTNSLMAANLSSNIFPSVRRITERAKKRITSSQSGIGVMP 402

Search completed: May 3, 2002, 11:02:16  
Job time: 51 sec

us-09-671-050-12.rag

Fri May 3 11:22:26 2002

\_\_\_\_\_



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:03:16 ; Search time 26.4 seconds  
(without alignments)  
908.901 Million cell updates/sec

Title: US-09-671-050-12

Perfect score: 1662  
Sequence: 1 MEKYELAKTGEISYGVFK.....RKARNEGRNRQQVLPKLS 315

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183.5	71.2	358	2 S23383	protein kinase (EC
2	1175.5	70.7	376	2 S22745	serine/threonine p
3	581	35.0	305	2 S23382	protein kinase (EC
4	574.5	34.6	292	2 S40021	protein kinase (EC
5	571.5	34.4	294	1 A40444	protein kinase (EC
6	571.5	34.4	294	2 B40444	protein kinase (EC
7	570.5	34.3	302	1 S50474	protein kinase (EC
8	568.5	34.2	294	2 S22440	protein kinase (EC
9	567	34.1	298	2 A41227	protein kinase (EC
10	565.5	34.0	297	2 A37871	protein kinase (EC
11	561.5	33.8	294	2 S23095	protein kinase (EC
12	559	33.6	302	1 A44349	protein kinase (EC
13	557	33.5	297	1 S24913	protein kinase (EC
14	556.5	33.5	294	2 T49271	CELL DIVISION CONT
15	556	33.5	297	2 A29559	protein kinase (EC
16	556	33.5	302	2 T45977	Cyclin-dependent k
17	556	33.5	302	2 B44349	protein kinase (EC
18	554	33.3	301	1 S19209	protein kinase (EC
19	553	33.3	298	1 A44878	protein kinase (EC
20	551.5	33.2	288	1 S36619	protein kinase (EC
21	551	33.2	288	1 S42566	protein kinase (EC
22	550.5	33.1	294	1 J02243	protein kinase (EC
23	550.5	33.1	294	1 S57928	protein kinase (EC
24	550.5	33.1	294	1 S42049	protein kinase (EC
25	547.5	32.9	303	1 S06011	protein kinase (EC
26	547.5	32.9	308	1 S53538	protein kinase (EC
27	546.5	32.9	297	1 S12009	protein kinase cdc
28	545	32.8	292	1 A46365	protein kinase (EC
29	544.5	32.8	291	2 S23386	protein kinase (EC

30	544	32.7	293	2 JE0374	cyclin-dependent k
31	543	32.7	292	1 I49592	protein kinase (EC
32	543	32.7	292	1 A45091	protein kinase (EC
33	543	32.7	297	1 A36074	protein kinase (EC
34	543	32.7	301	1 S42101	protein kinase (EC
35	541.5	32.6	294	1 S31332	protein kinase (EC
36	540	32.5	296	2 S24386	protein kinase PHO
37	539	32.4	302	1 OKBY85	protein kinase cdc
38	535.5	32.2	302	2 T17115	CDK-activating pro
39	530.5	31.9	346	2 A4820	protein kinase (EC
40	530	31.9	294	2 S51008	protein kinase (EC
41	528	31.8	292	2 S22441	protein kinase (EC
42	527.5	31.7	346	2 A56231	MOL5/CDK-activatin
43	527	31.7	346	1 I78840	protein kinase (EC
44	525	31.6	332	1 S41003	protein kinase (EC
45	525	31.6	346	1 I48157	protein kinase (EC

ALIGNMENTS

RESULT 1

S23383

protein kinase (EC 2.7.1.37) cdc2-related KKIALRE - human

C:Species: Homo sapiens (man)

C:Date: 15-Oct-1994 #sequence\_revision 15-Oct-1994 #text\_change 18-Jun-1999

C:Accession: S23383; S22744

R:Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.;

EMBO J. 11, 2909-2917, 1992

A:Title: A family of human cdc2-related protein kinases.

A:Reference number: S23382; MUID:92347325

A:Accession: S23383

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-358 <MEY>

A:Cross-references: EMBL:X66358; NID:g36614; PIDN:CAA47002.1; PID:g36615

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F;3-278/Domain: protein kinase homology <KIN>

F;11-19/Region: protein kinase ATP-binding motif

F;34,52,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 71.2%; Score 1183.5; DB 2; Length 358;  
Best Local Similarity 71.8%; Pred. No. 3.2e-49;  
Matches 211; Conservative 41; Mismatches 41; Indels 1; Gaps 1;

QY 1 MEKYELAKTGEISYGVFKCNKTSQGVAVKVFVESEDDPVVKIALREIRMLKOLKH 60

Db 2 MEKYEKIGKIGEGSYGVFKNRDGTQGVAVKVFLESEDDPVVKIALREIRMLKOLKH 61

QY 61 PNLVNLIEVFRKRKMLHVFECYDHTLLNELERNPNMGADVGVKSVLWQTLQALNFCIH 120

Db 62 PNLVNLIEVFRKRKMLHVFECYDHTLVHLDYRQGVPEHLVKSITWQTLQAVNFCCHK 121

QY 121 NCIHRIKPNELITKOGIKICDFGFAOILI-PGDAYTDYVATRWYRAPELVGDTQYG 179

Db 122 NCIHRIKPNELITKHSYKICDFGFAOILI-PGDAYTDYVATRWYRAPELVGDTQYG 181

QY 180 SSVDTWAICGVFAELLPGOPKGSVDVQDLYLIITLTKLIPRQSTFKNGFFHGIST 239

Db 182 PPVDVMAICGVFAELLPGOPKGSVDVQDLYLIITLTKLIPRQSTFKNGFFHGIST 241

QY 240 PEPEDMETIEKFSDDVHPVAMKGLKMPDRLTCSQLLESSYFDSFQDEAQ 293

Db 242 PDPEDMEPLKLFNYSYPALGILLKGLHMDPTERTLTCQELLHHPYFNIREIE 295

RESULT 2

S22745

serine/threonine protein kinase KKIALRE (EC 2.7.1.-) - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 19-Dec-1997

C:Accession: S22745  
R: Meyerson, M.L.  
submitted to the EMBL Data Library, May 1992  
A:Reference number: S22743  
A:Accession: S22745  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-376 <MEY>  
A:Cross-references: EMBL:X66359  
C:Genetics:  
A:Introns: 152/3; 170/3  
C:Superfamily: kinase-related transforming protein; protein kinase homology  
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F:3-296/Domain: protein kinase homology <KIN>  
F:11-19/Region: protein kinase ATP-binding motif

Query Match 70.7%; Score 1175.5; DB 2; Length 376;  
Best Local Similarity 67.6%; Pred. No. 7.9e-49;  
Matches 211; Conservative 41; Mismatches 41; Indels 19; Gaps 1;

QY 1 MEKYEKLAKTGEGSYGVFKCRNKTSGQVAVKFFVESEDDPVVKKIALREIRMLKQLKH 60  
DB 2 MEKYEKLGKIGEGSYGVFKCRNKTSGQVAVKFFVESEDDPVVKKIALREIRMLKQLKH 61  
QY 61 PNLVNLLEVERRRKRKMHVFEYCDHTLLNELRNPNVGADGVKSVLMOTLOALNFCIH 120  
DB 62 PNLVNLLEVERRRKRKMHVFEYCDHTVHLELDYRGVPEHLVKSITWGLQAVNFCCHK 121  
QY 121 NCIHRIKPNENILITKOGIIKICDGFQAQIL-IPGDAYTDYVATRWYRAPELLVGD 161  
DB 122 NCIHROVKPENILITKHSVIKLCDFGFAKLFXFOPSAAVCFPCSTGTGTPSYDYTDYV 181  
QY 162 ATWYRAPELLVGDYQYSSVDIWAIGCVFAELLTGQPLPGKSDVDQYLYLIRTKGLI 221  
DB 182 ATWYRSPPELLVGDYQYGPVDMVAIGCVFAELLSGVPLPGKSDVDQYLYLIRTKGLI 241  
QY 222 PRHOSIFKNGFFGHISIPPEDEMTELEKFSDVHPVALNFMKGCCKMNPDRITCSQLL 281  
DB 242 PRHQVFSTNQVFSGVKIPDPEDMEPLEKFNFTSYPALGLLAGCLMDPTERTCQULL 301  
QY 282 ESSYFDSFOEAQ 293  
DB 302 HHYFENIREIE 313

RESULT 3  
S23382  
protein kinase (EC 2.7.1.37) cdk3 - human  
C:Species: Homo sapiens (man)  
C:Date: 18-Jun-1993 #sequence\_revision 18-Jun-1993 #text\_change 18-Jun-1999  
C:Accession: S23382; S22743  
R: Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; Tsai  
EMBO J. 11, 2909-2917, 1992  
A:Title: A family of human cdc2-related transforming protein kinases.  
A:Reference number: S23382; MUID:92347325  
A:Accession: S23382  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-305 <MEY>  
A:Cross-references: EMBL:X66357; NID:g36612; PIDN:CAA47001.1; PID:g36613  
C:Genetics:  
A:Gene: GDB:CDK3  
A:Cross-references: GDB:283456  
A:Map position: 12q13-12q13  
C:Superfamily: kinase-related transforming protein; protein kinase homology  
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase  
F:2-255/Domain: protein kinase homology <KIN>  
F:10-18/Region: protein kinase ATP-binding motif  
F:33.51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 35.0%; Score 581; DB 2; Length 305;

Best Local Similarity 39.8%; Pred. No. 4.5e-21;  
Matches 123; Conservative 66; Mismatches 96; Indels 24; Gaps 8;  
QY 1 MEKYEKLAKTGEGSYGVFKCRNKTSGQVAVKFF-VESEDDPVVKKIALREIRMLKQ 57  
DB 1 MDMFQKVEKIGEGSYGVYKAKNRETQGLVALKIRLDEMEGVP---STAIRISLUKE 57  
QY 58 LKHPNLVNLLEVERRRKRKMHVFEYCDHTLLNELRNPNVGADGVKSVLMOTLOALNF 116  
DB 58 LKHPNLVNLLEVERRRKRKMHVFEYCDHTLLNELRNPNVGADGVKSVLMOTLOALNF 117  
QY 117 CHINCHTRDIRKPNENILITKOGIIKICDGFQAQIL-IPGDAYTDYVATRWYRAPELLVGD 175  
DB 118 CHSHRVTHRLDLPQNLLINELGAIKADFCARAGVPLRTYTHEVTVLWYRAPELLVGS 177  
QY 176 TQYSSVDIWAIGCVFAELLTGQPLPGKSDVDQYLYLIRTKGLIPRHOISTFKSNGFFH 235  
DB 178 KFYTTAVDIWSIGCIFAEWVTRKALPFGDSEIDQLFRIFRMJGT-----PSEDTPW 228  
QY 236 GIS-IPE-----PE-DMETLEKFSDVHPVALNFMKGCCKMNPDRITCSQLLESSYFDS 288  
DB 229 GVTQLPDYKSGSPKTRKGLIEIVPNLEPEGRDLIMQLLOYDPSQRITAKTALAHYPFSS 288  
QY 289 FOEAQIKRK 297  
DB 289 PEPSPAARQ 297  
RESULT 4  
S40021  
protein kinase (EC 2.7.1.37) cdc2 homolog - slime mold (Dictyostellium discoideum)  
C:Species: Dictyostellium discoideum  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 18-Jun-1999  
R: Michaelis, C.E.; Weeks, G.  
submitted to the EMBL Data Library, August 1992  
A:Description: The unicellular organism Dictyostellium discoideum possesses a highly r  
A:Reference number: S40021  
A:Accession: S40021  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-292 <MIC>  
A:Cross-references: EMBL:L00652; NID:g167695; PIDN:AAA16056.1; PID:g167696  
R: Michaelis, C.; Weeks, G.  
Biochim. Biophys. Acta 1179, 117-124, 1993  
A:Title: The isolation from a unicellular organism, Dictyostellium discoideum, of a hi  
A:Reference number: S39071; MUID:94032415  
A:Accession: S39071  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-18,'Y',20-188,'G',190-292 <MI2>  
A:Cross-references: EMBL:L00652  
C:Complex: In various organisms, cdc2 has been identified as a component of the M-pha  
C:Superfamily: kinase-related transforming protein; protein kinase homology  
C:Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threo  
F:2-254/Domain: protein kinase homology <KIN>  
F:10-18/Region: protein kinase ATP-binding motif  
F:33.51,126,128/Active site: Lys, Glu, Asp, Lys #status predicted  
Query Match 34.6%; Score 574.5; DB 2; Length 292;  
Best Local Similarity 37.5%; Pred. No. 8.8e-21;  
Matches 115; Conservative 64; Mismatches 93; Indels 35; Gaps 5;  
QY 1 MEKYEKLAKTGEGSYGVFKCRNKTSGQVAVKFFVESEDDPVVKKIALREIRMLKQLKH 60  
DB 1 MEKYSKIEKLGEGSYGVNKAKNREYGEIVALKRIRLDSDEGVPCPTAIRISLUKE 60  
QY 61 PNLVNLLEVERRRKRKMHVFEYCDHTLLNELRNPNVGADGVKSVLMOTLOALNFCIH 120  
DB 61 PNLVNLLEVERRRKRKMHVFEYCDHTLLNELRNPNVGADGVKSVLMOTLOALNFCIH 120  
QY 121 NCIHRIKPNENILITKOGIIKICDGFQAQIL-IPGDAYTDYVATRWYRAPELLVGDYQY 179



R:Colasanti, J.; Tyers, M.; Sundaresan, V.  
Proc. Natl. Acad. Sci. U.S.A. 88, 3377-3381, 1991  
A:Title: Isolation and characterization of cDNA clones encoding a functional p34 (cdc2)  
A:Reference number: A40444; MUID:91195354  
A:Accession: B40444  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-294 <COL>  
A:Cross-references: GB:M60526  
C:Superfamily: kinase-related transforming protein; protein kinase homology  
C:Keywords: ATP; phosphotransferase  
F:2-256/Domain: protein kinase homology <KIN>  
F:10-18/Region: protein kinase ATP-binding motif  
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.4%; Score 571.5; DB 2; Length 294;  
Best Local Similarity 37.6%; Pred. No. 1.2e-20;  
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGOVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60  
DB 1 MEQYKVEKIGEGTYGVVYKALDKTANETIALKIRLEQEDGEGVSPSTAIRISLLKEMNH 60  
QY 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLN-----ELERNPNGVADGVKSVLWQTLQAL 114  
DB 61 GNIVRLHDVHSEKRIYLVFEYLDLDLKKFMDSCPEFAKNPT-----LIKSYLYQLHGV 115  
QY 115 NFCHIHNCIHRDIKPNELITKQ-GIIKICDFGPAQL-IPGDATYDVATRWYRAPELL 172  
DB 116 AYCHSRVLRHLDLPQNLLIDRNALKLADGLARAFGIPVTRTFHEVTLWYRAPEIL 175  
QY 173 VGDYQYSSVDIATGCVFELLTGQPLWPGKSDVDQYLIITLGLKLIPIHQSIKFSNG 232  
DB 176 LGARQYSTPVDVWSVGGCIFAEMVNQKPLFPGDSEIDELFKIFRLVG--TPNEQ----- 227  
QY 233 PFHGIS-IP-----EPEDMETLEEKFSVDHPVALNFMKCLKMPDRLTCSOLLE 282  
DB 228 -WPGVSLPDKFTAFPRWQAQDLATI---VPLNLEPAGLDLLSKMLRYEFSKRITARQALE 283  
QY 283 SSYFDSFQEAQ 293  
DB 284 HEYFKDLEWVQ 294

RESULT 7  
150474  
protein kinase (EC 2.7.1.37) cdc2 [similarity] - goldfish  
C:Species: Carassius auratus (goldfish)  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
C:Accession: 150474  
R:Kajura, H.; Yamashita, M.; Katsu, Y.; Nagahama, Y.  
Dev. Growth Differ. 35, 647-654, 1993  
A:Title: Isolation and characterization of goldfish cdc2, a catalytic component of  
A:Reference number: 150474  
A:Accession: 150474  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-302 <KAJ>  
A:Cross-references: GB:D17758; NID:g471097; PIDN:BAA04605.1; PID:g471098  
C:Superfamily: kinase-related transforming protein; protein kinase homology  
C:Keywords: ATP; phosphotransferase  
F:2-256/Domain: protein kinase homology <KIN>  
F:10-18/Region: protein kinase ATP-binding motif  
F:33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.3%; Score 570.5; DB 1; Length 302;  
Best Local Similarity 38.6%; Pred. No. 1.4e-20;  
Matches 120; Conservative 68; Mismatches 90; Indels 33; Gaps 7;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGOVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59  
DB 1 MEQYKVEKIGEGTYGVVYKALDKTANETIALKIRLEQEDGEGVSPSTAIRISLLKEMNH 59

```
Db 1 MDDYLKIEKIGETGYGVYVYGRNKTGQVAMKKIRLSEBEG-VFSTAVREISLLKELQ 59
QY 60 HPNVLNVLIEVFRKRKMHVFEYCDHTLLNELERNPG--VADGVYKSVLWQTLQALNFC 117
Db 60 HPNVRLDLVMDKSLYLVFEFLSMDLKYLDSPSGQFMDPLMKVSLYQILGELIFC 119
QY 118 HINCIRHDKPENILITKOGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDIT 176
Db 120 HCRVLRDLKPQNLIDNKGVIKLADEGLARAFGVPVRYTHEVVTWLYRAPEVLGAS 179
QY 177 QYSSVDIWAICGVFAELLTGQPLWPKGSVDQVLYLIRTLG-----KLIPRHQS 226
Db 180 RYSTPVDVMSIGTIFAEALATKKPLFGHDSIDQLFRIFRTLGTNNNEVMPDVESLPDYN 239
QY 227 IF---KSNQFFHGISIPEDMETLEKFSVDHPVALFMKGLCKMKNPDRLTCSOLLES 283
Db 240 TFPKWSG-----NLASTVKNLKNKIDILLTKMLIYDPPKRISARQAMTH 284
QY 284 SYFDSFQEAQI 294
Db 285 PYFDDLDKSTL 295

RESULT 8
S22440
protein kinase (EC 2.7.1.37) cdc2 homolog 1 - rice
C:Species: Oryza sativa (rice)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S22440
R:Hashimoto, J.; Hirabayashi, T.; Hayano, Y.; Hata, S.; Ohashi, Y.; Suzuka, I.; Utsugi, M. Genet. 233:10-16, 1992
A:Title: Isolation and characterization of cDNA clones encoding cdc2 homologues from Ory
A:Reference number: S22440; MUID:94293101
A:Accession: S22440
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <HNS>
A:Cross-references: EMBL:X60374; NID:g20342; PIDN:CAA42922.1; PID:g20343
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k
F:2-356/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:15/Binding site: phosphate (Tyr) (covalent) #status predicted
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.2%; Score 568.5; DB 2; Length 294;
Best Local Similarity 37.6%; Pred. No. 1.7e-20;
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;
QY 1 MEKYEKATGEGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIMLKQLKH 60
Db 1 MEQYEKEKIGETGYGVYVYGRNKTGQVAMKKIRLSEBEG-VFSTAVREISLLKELH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPGVADGVYKSVLWQTLQAL 114
Db 61 GNIVRLHDVTHSEKRIYLVFEYLDLQKFKMDSCPEFAKNPT-----LIKSYLYQILRGV 115
QY 115 NFCHINCHIRDKPENILITKQ-GTIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELL 172
Db 116 AYCHSHRVLRLDKPQNLIDRNTNALKLADFGARAFGVPVRYTHEVVTWLYRAPELL 175
QY 173 VGDQYGVSSVDIWAICGVFAELLTGQPLWPKGSVDQVLYLIRTLGKLIPRHQSIFKNSG 232
Db 176 LGSQYSTPVDVMSVGGCIFAEAMVKNQKPLFGDSEIDELFKIFRVLG--TNEQS----- 227
QY 233 FFHGT-SIP-----EPEDMETLEKFSVDHPVALFMKGLCKMKNPDRLTCSOLLE 282
Db 228 -WPGVSSLPDYKSAFPKWAQDLATI---VPTLDPAGLDLSLKMRLYEPNKRITARQALE 283
QY 283 SSYFDSFQEAQI 293
Db 284 HEYFKDLEMVQ 294
```

```
RESULT 9
A41227
protein kinase (EC 2.7.1.37) cdk2 - human
N:Alternate names: Egl homolog; protein kinase p34
C:Species: Homo sapiens (man)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
C:Accession: A41227; S17873; S16520
R:Ninomiya-Tsujii, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.
Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991
A:Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of
A:Reference number: A41227; MUID:92020980
A:Accession: A41227
A:Molecule type: mRNA
A:Residues: 1-298 <NIN>
A:Cross-references: GB:M68520; NID:g180177; PIDN:AAA35667.1; PID:g180178
R:Tsal, L.H.; Haylow, E.; Meyerson, M.
Nature 353, 174-177, 1991
A:Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E
A:Reference number: S17873; MUID:91367262
A:Accession: S17873
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-298 <TSA>
A:Cross-references: GB:X62071; NID:g312802; PIDN:CAA43985.1; PID:g312803
R:Elledge, S.J.; Spottswood, M.R.
EMBO J. 10, 2653-2659, 1991
A:Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc
A:Reference number: S16520; MUID:91330891
A:Accession: S16520
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-176, 5', 178-208 <ELL>
A:Cross-references: EMBL:X61622; NID:g29848; PIDN:CAA43807.1; PID:g29849
C:Genetics:
A:Gene: GDB:CDK2
A:Cross-references: GDB:128984; OMIM:116953
A:Map position: 12q13-12q13
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; ser
F:2-255/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif
F:14,160/Binding site: phosphate (Thr) (covalent) #status predicted
F:15/Binding site: phosphate (Tyr) (covalent) #status predicted
F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 34.1%; Score 567; DB 2; Length 298;
Best Local Similarity 40.0%; Pred. No. 2e-20;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;
QY 1 MEKYEKATGEGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIMLKQLKH 60
Db 1 MENFOKEKIGETGYGVYKARNKLGENVALKIRLDTETEGVSTAIRLSLLKELAH 60
QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNP-NGVADGVYKSVLWQTLQALNFC 119
Db 61 PNIVRLDLVHTENKLYLVFEFLHQDLKFKMDASALTG1PLPLIKSYLFQLOGIAPCHS 120
QY 120 HNCIRHDKPENILITKOGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDIT 178
Db 121 HVRVLRDLKPQNLIDNKGVIKLADEGLARAFGVPVRYTHEVVTWLYRAPEILLGCKY 180
QY 179 GSSVDIWAICGVFAELLTGQPLWPKGSVDQVLYLIRTLGKLIPRHQSIFKNSGFFHGT- 237
Db 181 STAVDLSGLGICFAEAVTRALFPDSEIDQLFRIFRTLTGT-----PDEVVWFGVT 231
QY 238 STPE--PDMETLEKFSVDHPV---ALNFMKGLCKMKNPDRLTCSOLLESFF 286
Db 232 SMPDYKSPFKWARQDFSKVVPDLEDDGRSLLSQMLHYDPNKRISAKAALAHPEF 286

RESULT 10
```





Fri May 3 11:22:28 2002

Db 181 YSTPVDIWSGCIFAEMISOKFLFPDSEIDQLKIFRIMGTPTPYEDTWRCVTSLPDYKSA 240  
Qy 228 FKSNGFFHGISIPEDMETLEKFSOVHPVALNFMKGLKMNPDRLTCSQLLESSYF 286  
Db 241 FPK-----WKPTDLETF---VPNLDPGVDLLSKMLMDPTKRINARAALAEHEYF 287

RESULT 15  
A29539  
protein kinase (EC 2.7.1.37) cdc2 - human  
N:Alternate names: cell division control protein 2 (CDC2)  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 21-Jul-2000  
C:Accession: A29539 P.  
R:Lee, M.G.; Nurse, P.  
Nature 327, 31-35, 1987  
A:Title: Complementation used to clone a human homologue of the fission yeast cell cycle  
A:Reference number: A29539; MUID:87201915  
A:Accession: A29539  
A:Molecule type: mRNA  
A:Residues: 1-297 <LEE>  
A:Cross-references: GB:X05360; NID:g29838; PIDN:CAA28963.1; PID:g29839  
C:Genetics  
A:Gene: GDB:CDC2  
A:Cross-references: GDB:119052; OMIM:116940  
A:Map position: 10q21.1-10q21.1  
C:Superfamily: kinase-related transforming protein; protein kinase homology  
C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine  
F:2-256/Domain: protein kinase homology <Kin>  
F:10-18/Region: protein kinase ATP-binding motif  
F:14.161/Binding site: phosphate (Thr) (covalent) #status predicted  
F:15/Binding site: phosphate (Tyr) (covalent) #status predicted  
F:33.51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 33.58; Score 556; DB 2; Length 297;  
Best Local Similarity 39.48; Pred. No. 6,6e-20;  
Matches 123; Conservative 61; Mismatches 96; Indels 32; Gaps 8;

Qy 1 MEKYEKLAKTGEGSYGVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKOLK 59  
Db 1 MEDYTKIEKIGETGYGVYKGRHKTTGQVAVAMKKIRLESEEG-VPSTAIRLSLLKELR 59

Qy 60 HNLVNLVEYFRKRKMHLVFEYCDHTLLNELERNPENG--VADGVKISVLMTQLQALNFC 117  
Db 60 HNRVLSLDVLMQDSRLYLIFEFLSMDLKKYLDSPPPGOYMDSSLVKSYLYQLQIGIVFC 119

Qy 118 HTHNCIHRDIKPNILITKQGIKICDFEQAAIL-IPGDAYTDVYATWYRAPELLVGDT 176  
Db 120 HSRRLVRLDLKPQLLDDKGTIKLADEGLARAFGIPRYTHVVTILWYRSPEVLGSA 179

Qy 177 QYGSVDIWAIGCVFAELLTGQPLMPGKSDVDQLYLIIRTGLKLIPRHQSTFKSNGFFHG 236  
Db 180 RYSTPVDIWSGTFIABLATKPLFGDSEIDQLFRIFRALG--TPNNE----- 226

Qy 237 ISIPEDMETLEKFSOVHPVA-----LNMKGLKMNPDRLTCSQLLESS 284  
Db 227 -VMPEVSQDYKNTFFPKWPGSLASHVKNLDENGLDLSKMLLYDPAKRISGKMAINHP 285

Qy 285 YFDSFOEAQIKR 296  
Db 286 YFNDL-DNQIKK 296

Search completed: May 3, 2002, 11:03:17  
Job time: 111 sec



Result No.	Query			ID	Description
	Score	Match	Length		
1	1183.5	71.2	358	1	KKTA_HUMAN
2	720.5	43.4	1030	1	STK3_HUMAN
3	581	35.0	305	1	CDK3_HUMAN
4	574.5	34.6	292	1	CDK3_DICDI
5	571.5	34.4	297	1	CDK2_XENLA
6	570.5	34.3	302	1	CD2_CARAU
7	568.5	34.2	294	1	CC21_ORYSA
8	567	34.1	298	1	CDK2_HUMAN
9	566.5	34.1	294	1	CC2_MATZE
10	563	33.9	298	1	CDK2_CRIGR
11	561.5	33.8	294	1	CD2A_ARATH
12	561	33.8	298	1	CDK2_RAT
13	559	33.6	298	1	CDK2_MESAU
14	559	33.6	302	1	CC21_XENLA
15	557	33.5	297	1	CC2_RAT
16	556	33.5	297	1	CC2_BOVIN
17	556	33.5	297	1	CC2_HUMAN
18	556	33.5	302	1	CC22_XENLA
19	554	33.3	301	1	CC21_TRYBB
20	553	33.3	298	1	CDK2_CARAU
21	551.5	33.2	294	1	CC2_VIGUN
22	551.5	33.2	311	1	CC22_TRYBB
23	551	33.2	288	1	CC2H_PLAFK
24	548	33.0	297	1	CC2_MOUSE
25	547.5	32.9	303	1	CC2_CHICK
26	546.5	32.9	297	1	CC2_DROME
27	545	32.8	292	1	CDK5_BOVIN
28	545	32.8	292	1	CDK5_RAT
29	544	32.7	292	1	CDK5_HUMAN
30	543	32.7	292	1	CDK5_MOUSE
31	543	32.7	301	1	CC21_TRYGO
32	541.5	32.6	294	1	CC22_MESAU
33	540	32.5	296	1	CC2_DICDI

```
Matches 211; Conservative 41; Mismatches 41; Indels 1; Gaps 1;
QY 1 MEKYEKLAKTGGSGYGVVFCRNKTSQGVAVKVFSEDDPVVKKIALREIRMLKOLKH 60
Db 2 MEKYEKIGKIGEGSYGVVFCRNKTDGQIVAKKFLSEDDPVVKKIALREIRMLKOLKH 61
QY 61 PNLVNLIEVFRKKRMLHVFYCDHTLLNELRNPNVGVADGVKSVLWQTLQALNFCCHI 120
Db 62 PNLVNLIEVFRKKRMLHVFYCDHTLVHLDYQGVPEHUVKSVTWQTLQAVNFCCHK 121
QY 121 NCHRDIKPENILITKOGIKICDFGFAQILI-PGDAYDYVATRWYRAPELLVGDGTQV 179
Db 122 NCHRDVAPENILITKHSVVKICDFGFARLITGSDYDYVATRWYRSPPELLVGDGTQV 181
QY 180 SSVDIWATGCVFALLTGQPLWPGKSDVDQLYLITLGLKLPKHSITKSNFGFHGISI 239
Db 182 PPDVWATGCVFALLTGQPLWPGKSDVDQLYLITLGLKLPKHSITKSNFGFHGISI 239
QY 240 PEPEDMETLEEKFSQVHPVAFNFMKGLKMPDDRLTCSOLLESSYFDSFQEAQ 293
Db 242 PPDPEMEPLEKFPNISYPALGGLKGLHMDPTERTCQLLHHPYFENIRETE 295
RESULT 2
STR9_HUMAN
ID STK9_HUMAN STANDARD: PRT; 1030 AA.
AC 076039; Q14198;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE SERINE/THREONINE-PROTEIN KINASE 9 (EC 2.7.1.37).
GN STK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=98389628; PubMed=9721213;
RA Montini E., Andolfi G., Caruso A., Buchner G., Walpole S.M.,
RA Mariani M., Consalez G.G., Trump D., Ballabio A., Franco B.;
RT "Identification and characterization of a novel serine-threonine
RL kinase gene from the Xp22 region."
RN 2
RP SEQUENCE OF 339-789 FROM N.A.
RA Krause S.W., Rehli M., Kreutz M., Schwarzfischer L., Paulauskis J.D.,
RA Andreesen J.D.;
RT "Differential screening leads to novel genetic markers of monocyte
RL to macrophage maturation."
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
CC -1- ENZYME REGULATION: COULD BE ACTIVATED BY PHOSPHORYLATION ON
CC TYROSINE AND THREONINE. PROBABLY BOTH PHOSPHORYLATIONS ARE
CC REQUIRED FOR ACTIVITY.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG, KIDNEY, PROSTATE,
CC OVARY, PLACENTA, PANCREAS AND TESTIS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 415.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Y15057; CAA75342.1;
DR EMBL; X89059; CAA61445.1; ALT_FRAME.
```

```
DR MM; 300203;
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 13 297 PROTEIN KINASE.
FT NP_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 135 135 BY SIMILARITY.
FT MOD_RES 169 169 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 171 171 (POTENTIAL).
FT MOD_RES 171 171 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT DOMAIN 784 789 (POTENTIAL).
FT CONFLICT 339 340 HR -> GT (IN REF. 2).
FT CONFLICT 541 541 L -> W (IN REF. 2).
FT CONFLICT 731 764 MISSING (IN REF. 2).
SQ SEQUENCE 1030 AA; 115537 MW; 8A1C9C438610FF08 CRC64;
Query Match 43.4%; Score 720.5; DB 1; Length 1030;
Best Local Similarity 44.4%; Pred. NO. 9.7e-45;
Matches 139; Conservative 69; Mismatches 98; Indels 7; Gaps 4;
QY 1 MEKYEKLAKTGGSGYGVVFCRNKTSQGVAVKVFSEDDPVVKKIALREIRMLKOLKH 60
Db 10 MNKFETLGVVGGAGVVLKCRUKETHEIVALKFKDSENEVEYKETTRELKMLRTLKQ 69
QY 61 PNLVNLIEVFRKKRMLHVFYCDHTLLNELRNPNVGVADGVKSVLWQTLQALNFCCHI 120
Db 70 ENIVELKEAFRRCKLYLVEFVEYKMKLLELLEMPNGVPEKYSYIYQLIKAIHWCHKN 129
QY 121 NCHRDIKPENILITKOGIKICDFGFAQILIPGD--AYTDVATRWYRAPELLVGDGTQY 178
Db 130 DIVHRDIKPENILLSHNDVLKLCDFGFARNLSEGNANNTVEYATRWYRSPPELLLG-APY 188
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLITLGLKLPKHSITKSNFGFHGIS 238
Db 189 GKSDVMSVGCILGELSDGQPLFGESIDQLEFTIOKVLPGLPSQMKLFYSNPFHGLR 248
QY 239 IPEPEDEMETLEEKFSQV-HPVALNFMKCLMKMPDRLTCSOLLESSYFDSFQEAQIKRK 297
Db 249 FPAVNHPOSLEERYILNSVLIDLMKNLKLDPADRYLTQCLN--HPTFQTQRLDDR 305
QY 298 ARNEGRNRRQQV 310
Db 306 SPSRSAKRKPVH 318
RESULT 3
CDK3_HUMAN
ID CDK3_HUMAN STANDARD: PRT; 305 AA.
AC Q00526;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 3 (EC 2.7.1.1-).
GN CDK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=fetal brain;
RX MEDLINE=92347325; PubMed=1639063;
RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,
RA Harlow E., Tsai L.-H.;
```



"A family of human cdc2-related protein kinases.";  
EMBO J. 11:2909-2917(1992).  
-1- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.  
INTERACTS WITH A YET UNKNOWN TYPE OF CYCLIN. CAN PHOSPHORYLATE HISTONE H1.  
-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CDC2/CDKX SUBFAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
-----  
EMBL: X66357; CAA47001.1; -;  
PIR: S22743; S22743.  
PIR: S23382; S23382.  
HSSP: P24941; IAQ1.  
MIM: 123828; -;  
InterPro: IPR000719; Euk\_pkinase.  
InterPro: IPR002290; Ser\_thr\_kin\_actsite.  
Pfam: PF00069; pkinase; 1.  
SMART: SM00220; S\_TKc; 1.  
PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
Transferase: Serine/threonine-protein kinase; ATP-binding;  
Cell cycle; Cell division; Mitosis; Phosphorylation.  
DOMAIN 4 286 PROTEIN KINASE  
NP\_BIND 10 18 ATP (BY SIMILARITY).  
BINDING 33 33 ATP (BY SIMILARITY).  
ACT\_SITE 127 127 BY SIMILARITY.  
FT ACT\_SITE 127 127 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 160 160 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 305 AA; 35045 MW; 1128BE096EA262A CRC64;  
SEQUENCE 305 AA; 35045 MW; 1128BE096EA262A CRC64;  
-----  
Query Match 35.0%; Score 581; DB 1; Length 305;  
Best Local Similarity 39.8%; Pred. No. 3e-35; Indels 24; Gaps 8;  
Matches 123; Conservative 66; Mismatches 96;  
QY 1 MEKYEKLAKTEGSGYGVVFKCRNKTSGQVAVKFF--VESEDDPVVKKIALREIRMLKQ 57  
DB 1 MDMFQVKEIGEGTYGVVYKAKNRETGQVAVKIRLDLEMEGVP---STAIREISLLKE 57  
QY 58 LKHPNLNIEVFRKRKMLVFEYCDHTLLNELERNPNC-VADGVKSVLWOTLQALNF 116  
DB 58 LKHPNIVRLDDVHNERKYLVEFLSQDLKMYDSTPGSELPLHLIKSYLFQLQGVSF 117  
QY 117 CHTHNCIRDKIPENILITKOGIIKICDGFQAQL-IPGDAYTDYVATRWYRAPELLVGD 175  
DB 118 CHSHRVHDLKPNQLLNLGALKAIDFLARAFGVPLRTYTHEVYTLWYRAPELLLGS 177  
QY 176 TQGVSSVDVWAGCVFELLTGOPLWPKGSDVDOLYLIITLGLKIPRHQSIFKSGFFH 235  
DB 178 KFYTTAVDILWSIGCFADWIRKALFPDGEIDQLFRIFRMIGT-----PSEDTPW 228  
QY 236 GIS-IPE-----PE-DMETLEEFSDVHPVALNFMKGLKMNPPDRLTCSQLLESFDS 288  
DB 229 GVTQLDPDGSPKPKRGLLEIVPNLEPEGRDLMLQLQYDPSQIRAKTALAHPYFSS 288  
QY 289 FOEAIAKRR 297  
DB 289 PEPSAARQ 297  
RESULT 4  
CC2H\_DICDI  
ID CC2H\_DICDI  
AC P34117;

01-FEB-1994 (Rel. 28, Created)  
01-FEB-1994 (Rel. 28, Last sequence update)  
15-JUL-1999 (Rel. 38, Last annotation update)  
CDC2-LIKE SERINE/THREONINE-PROTEIN KINASE CRP (EC 2.7.1.-).  
CRPA OR CRP.  
Dictyostelium discoideum (Slime mold).  
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
NCBI\_TaxID=44689;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=94032415; PubMed=8218353;  
McClellis C.E.; Weeks G.;  
"The isolation from a unicellular organism, Dictyostelium discoideum, of a highly-related cdc2 gene with characteristics of the PCTAIRE subfamily.";  
Biochim. Biophys. Acta 1179:117-124(1993).  
-1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES THE ENZYME (BY SIMILARITY).  
-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CDC2/CDKX SUBFAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
-----  
EMBL: L00652; AAAL6056.1; -;  
PIR: S40021; S40021.  
HSSP: P24941; IAQ1.  
Dictyob; DD05039; crpa.  
InterPro: IPR000719; Euk\_pkinase.  
InterPro: IPR002290; Ser\_thr\_kin\_actsite.  
Pfam: PF00069; pkinase; 1.  
SMART: SM00220; S\_TKc; 1.  
PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
Transferase: Serine/threonine-protein kinase; ATP-binding;  
Phosphorylation.  
DOMAIN 4 285 PROTEIN KINASE.  
NP\_BIND 10 18 ATP (BY SIMILARITY).  
BINDING 33 33 ATP (BY SIMILARITY).  
ACT\_SITE 126 126 BY SIMILARITY.  
FT ACT\_SITE 126 126 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 292 AA; 33226 MW; 361AB54C4E3BD41E CRC64;  
SEQUENCE 292 AA; 33226 MW; 361AB54C4E3BD41E CRC64;  
-----  
Query Match 34.6%; Score 574.5; DB 1; Length 292;  
Best Local Similarity 37.5%; Pred. No. 8.4e-35;  
Matches 115; Conservative 64; Mismatches 93; Indels 35; Gaps 5;  
QY 1 MEKYEKLAKTEGSGYGVVFKCRNKTSGQVAVKFFVESEDDPVVKKIALREIRMLKOLKH 60  
DB 1 MEKYSIKELGEGTYGVVYKAKNRETGQVAVKIRLDSEDEGVPCPTAIREISLLKELKH 60  
QY 61 PNVLNIEVFRKRKMLVFEYCDHTLLNELERNPNCVADGVKSVLWOTLQALNFCIH 120  
DB 61 PNIVRLHDVHTERKTLVFEYLDQDLKYLDECGEISKPTIKSFYOLLKGVAFCHDH 120  
QY 121 NCHRIDIKPENILITKOGIIKICDGFQAQL-IPGDAYTDYVATRWYRAPELLVGDYQY 179  
DB 121 RVLRDLKPNQLLNLGALKAIDFLARAFGVPLRTYTHEVYTLWYRAPELLVGDYQY 180  
QY 180 SSVDIWAGCVFAELLTGOPLWPKGSDVDOLYLIITLGLKIPRHQSIF- 228  
DB 181 TPIDISALCIAFAWASGRPLFGSGTSDQLFRIFKILGTNPEWSMPSITELPEYKTDGP 240  
QY 229 -----KSGNFFHGISIPEPEDMETLEKFSVHPVALNFMKGLKMNPPDRLTCSQLLES 283  
DB 229 -----KSGNFFHGISIPEPEDMETLEKFSVHPVALNFMKGLKMNPPDRLTCSQLLES 283

Db 241 VPEAHLSSIVHG-----LDEK-----GLNLKMLQYDPNORITAAALKH 282

Qy 284 SYEDSFQ 290

Db 283 PYDGL 289

RESULT 5

CDK2\_XENLA

ID CDK2\_XENLA STANDARD; PRT; 297 AA.

AC P23437

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-) (CDC2 HOMOLOG EGI PROTEIN KINASE).

DE EGI.

GN EGI.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus.

OC NCBI\_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;

RX MEDLINE=91126051; PubMed=1704128;

RA Paris J., le Guellec R., Couturier A., le Guellec K., Omilli F., Camonis J., Macneill S., Philippe M.;

RT "Cloning by differential screening of a Xenopus cDNA coding for a protein highly homologous to cdc2.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:1039-1043(1991).

RN [2]

RP PHOSPHORYLATION AT THR-160.

RX MEDLINE=93345457; PubMed=8393783;

RA Poon R.Y.C., Yamashita K., Adamczewski J.P., Hunt T., Shuttleworth J.;

RT "The cdc2-related protein p40M015 is the catalytic subunit of a protein kinase that can activate p33cdc2 and p34cdc2.";

RL EMBO J. 12:3123-3132(1993).

CC -1- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.

CC INTERACTS WITH CYCLINS A, D, OR E.

CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY SIMILARITY).

CC -1- DEVELOPMENTAL STAGE: SYNTHESIZED IN UNFERTILIZED EGG, BUT NO LONGER MADE IN THE EARLY EMBRYO.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC CDC2/CDKX SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X14227; CAA32443.1; .

DR PIR; A37871; A37871.

DR HSP; P24941; ICHK.

DR InterPro; IPR000719; Euk\_pkinase.

DR InterPro; IPR002290; Ser\_thr\_kin\_actsite.

DR Pfam; PF00069; pkinase; 1.

DR SMART; SM00220; S\_TKc; 1.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.

KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle; Cell division; Mitosis; Phosphorylation.

FT DOMAIN 4 286

FT NP\_BIND 10 18

FT BINDING 33 33

FT ACT\_SITE 127 127

FT MOD\_RES 14 14

PHOSPHORYLATION (BY SIMILARITY).

FT MOD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).

FT MOD\_RES 160 160 PHOSPHORYLATION (BY CAK).

SQ SEQUENCE 297 AA; 33969 MW; EC30204FCB8D198C CRC64;

Query Match 34.4%; Score 571.5; DB 1; Length 297;

Best Local Similarity 38.1%; Pred. No. 1.4e-34;

Matches 117; Conservative 65; Mismatches 84; Indels 41; Gaps 6;

QY 1 MEKYEKLAKTGGSGVGVFKRNTSGOVAVKFKVESEDDPVVKKTALREIMLKOLKH 60

DB 1 MENFQKVEKIGCTGVVYKARNRETGEIVALKKIRLDTETGVPSTAIRLSLJLH 60

QY 61 PNLVNLVFRKRRKHLVFEYCDHTLLNELER-NPNCVADGVTKSVLMOTLQALNFC 119

DB 61 PNIVKLLDVHTENKYLIVFEFLNQDLKKFMDRSNIGISLALVKSYLEQLQLGLAF 120

QY 120 HNCIHRDIPENILITKOGIHKICDFGAQIL-IPGDAYTDYVATRYRAPELLVGD 178

DB 121 HRVLHDLKPKQLLINSGLAKLADEGLARAFGVPRFTTHEVVTLMYRAPELLIG 180

QY 179 GSSVDIWAIGCVFAELLTGOPLWPKSDVDOLYLIIRTLG-----KLIPRH 228

DB 181 STAVDIWSLGCIFAEIMTRALPFGSEIDQLFRITRLGTDPDEVSWSGVTTPDY 240

QY 229 KSNPFPHGISIPEDMETLEEFKFSVHPVAFNFMKGLKMNPDRLTCSQLLESSY 288

DB 241 P-----KWIRODFSKVWP-----PLDEGDRLLAQMQLQ---YDS 271

QY 289 FQEAQIK 295

DB 272 NKIRISAK 278

RESULT 6

CC2\_CARAU

ID CC2\_CARAU STANDARD; PRT; 302 AA.

AC P51958;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG (EC 2.7.1.-) (P34 PROTEIN KINASE) (CYCLIN-DEPENDENT KINASE 1) (CDK1).

DE CDC2.

GN CDC2.

OS Carassius auratus (Goldfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Osteiophysi; Cypriniformes; Cyprinidae; Cyprininae; Carassius.

OX NCBI\_TaxID=7957;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Oocyte;

RA Kajihara H., Yamashita M., Katsu Y., Nagahama Y.;

RT "Isolation and characterization of goldfish cdc2, a catalytic component of maturation-promoting factor.";

RL Dev. Growth Differ. 35:647-654(1993).

CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC COMPONENT OF MPF.

CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY SIMILARITY).

CC -1- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN MATURE OOCYTES.

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC CDC2/CDKX SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----

DR EMBL: D17758; BAA04605.1; -.  
DR HSP: P24941; LHCK.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_kin\_actsite.  
DR Pfam: PF00069; pkinase; 1.  
DR SMART: SM00220; S\_TKC; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase: Serine/threonine-protein kinase; ATP-binding;  
KW Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.  
FT DOMAIN 4 287 PROTEIN KINASE.  
FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
FT BINDING 33 33 ATP (BY SIMILARITY).  
FT ACT\_SITE 128 128 BY SIMILARITY.  
FT MOD\_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).  
SQ SEQUENCE 302 AA; 34499 MW; 58DB812E19B311F5 CRC64;

Query Match 34.3%; Score 570.5; DB 1; Length 302;  
Best Local Similarity 38.6%; Pred. No. 1.7e-34;  
Matches 120; Conservative 68; Mismatches 90; Indels 33; Gaps 7;

QY 1 MEKYEKLAKTGEQGVVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIMLKOLK 59  
Db 1 MDDYLKTEKIGEGTYGVVYKGRNKTSGQVAVKKFLESEEG-VPSTAVREISLLKELQ 59  
QY 60 HNLVNLIEVFRKRKMHVPEYCDHTLLNELRNPNQ--VADGVKSVLMQTLQALNFC 117  
Db 60 HENVRLDLVMESEKLYLVEFLSMDLKKYLDSPSGFMDPMLKSVLYLQLEILGF 119  
QY 118 HTHNCIHHRDKPENILITKQGIKICDFGAQIL-IPGDVTDVATRWYRAPELIVGDT 176  
Db 120 HCRVRLHRLDKPQNLLIDNKGVIKADFGARAFGVPVRYVTHEVTLVRAPEVLGAS 179  
QY 177 QYGVSSVDIWAIGCVFAELITGQPLWPKGSDVDQLYLIITLQ-----KLIPRHOS 226  
Db 180 RYSTPDVMSIGTIFAELATKKPLPHGDSIDQLFRIFRTLTGTPNNEVWPDVSLPKYN 239  
QY 227 IF---KSNQFPHGISPEPEDMETLEEKFSVHPVALNFMKGLKMNPDRLTCSQLLES 283  
Db 240 TFPKWKSG-----NLASTVKMLDKNGIDILLTKMLIYDPPKRISARQAMTH 284  
QY 284 SYFDSFQEAQI 294  
Db 285 PYFDDLDKSTL 295

RESULT 7  
CC21\_ORYSA STANDARD; PRT; 294 AA.  
AC P29618;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG 1 (EC 2.7.1.-).  
GN CD22-1.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV\_NIPPONBARE;  
RX MEDLINE=92293101; PubMed=1376401;

RA Hashimoto J., Hirabayashi T., Hayano Y., Hata S., Ohashi Y.,  
RA Suzuki I., Utsugi T., Toh-E A., Kikuchi Y.;  
RT "Isolation and characterization of cDNA clones encoding cdc2  
RT homologues from Oryza sativa: a functional homologue and cognate  
RT variants";  
RL Mol. Gen. Genet. 233:10-16(1992).  
CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC  
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES  
CC THE REPEITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.  
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES  
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY  
CC SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC CDC2/CDCX SUBFAMILY.  
CC -----  
CC this SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: X60374; CAA42922.1; -.  
DR PIR: S22440; S22440.  
DR HSP: P24941; LHCK.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_kin\_actsite.  
DR Pfam: PF00069; pkinase; 1.  
DR SMART: SM00220; S\_TKC; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase: Serine/threonine-protein kinase; ATP-binding;  
KW Cell cycle; Cell division; Mitosis; Phosphorylation.  
FT DOMAIN 4 287 PROTEIN KINASE.  
FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
FT BINDING 33 33 ATP (BY SIMILARITY).  
FT ACT\_SITE 127 127 BY SIMILARITY.  
FT MOD\_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 15 15 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).  
FT MOD\_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).  
SQ SEQUENCE 294 AA; 34071 MW; 51322D93AEF4C131 CRC64;

Query Match 34.2%; Score 568.5; DB 1; Length 294;  
Best Local Similarity 37.6%; Pred. No. 2.3e-34;  
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYEKLAKTGEQGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIMLKOLKH 60  
Db 1 MEQYEKEEKIGEGTYGVVYKGRNKTSGQVAVKKFVESEDDPVVKKIALREIMLKOLKH 60  
QY 61 PNLVNLIEVFRKRKMHVPEYCDHTLLN-----ELERNPNGVADGVKSVLMQTLQAL 114  
Db 61 GNIVRLHDVYHSEKRIYLYVEFLDLDLKKEMDSCPEFAKNPT-----LIKSYLYOILRGV 115  
QY 115 NFCHTHNCIHHRDKPENILITKQ-GLIKICDFGAQIL-IPGDVTDVATRWYRAPEL 172  
Db 116 AYCHSHVRLHRLDKPQNLLIDRNTAKLADFGARAFGIPVRYVTHEVTLVRAPEL 175  
QY 173 VGTQYGVSSVDIWAIGCVFAELITGQPLWPKGSDVDQLYLIITLQKLIIPHQSIFKSG 232  
Db 176 LGSQRYSTPDVMSVGCIFAEVMYVQKPLPPGDSIDELFKIFRVLG--TPNEQS----- 227  
QY 233 FFIGI-SIP-----EPEDMETLEEKFSVHPVALNFMKGLKMNPDRLTCSQLE 282  
Db 228 -WPSLSLDPYKSAFPKQWQADLATI---VPTLPAGLDLSSKMLRYEPNKRITARQALE 283  
QY 283 SSYFDSFQEAQ 293  
Db 284 HEYFKDLEMQV 294

```

RESULT 8
ID CDK2_HUMAN STANDARD; PRT; 298 AA.
AC P24941;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-) (P33 PROTEIN KINASE).
OS CDK2.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CG2A.
RX MEDLINE=96313126; PubMed=8756328;
RA Russo A.A., Jeffrey P.D., Pavletich N.P.;
RT "Structural basis of cyclin-dependent kinase activation by
RT phosphorylation.";
RL Nat. Struct. Biol. 3:696-700(1996).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=97075215; PubMed=8917641;
RA Schulze-Gahmen U., de Bondt H.L., Kim S.-H.;
RT "High-resolution crystal structures of human cyclin-dependent kinase
RT 2 with and without ATP: bound waters and natural ligand as guides for
RT inhibitor design.";
RL J. Med. Chem. 39:4540-4546(1996).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=97475219; PubMed=9334743;
RA Lawrie A.M., Noble M.E.M., Tunnah P., Brown N.R., Johnson L.N.,
RA Endicott J.A.;
RT "Protein kinase inhibition by staurosporine revealed in details of
RT the molecular interaction with CDK2.";
RL Nat. Struct. Biol. 4:796-801(1997).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CGKSL.
RX MEDLINE=96182647; PubMed=8601310;
RA Bourne Y., Watson M.H., Hickey M.J., Holmes W., Rocque W., Reed S.I.,
RA Tainer J.A.;
RT "Crystal structure and mutational analysis of the human CDK2 kinase
RT complex with cell cycle-regulatory protein CksHsl.";
RL Cell 84:863-874(1996).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=98342369; PubMed=9677190;
RA Gray N.S., Wodicka L., Thunnissen A.-M.W.H., Norman T.C., Kwon S.,
RA Espinoza F.H., Morgan D.O., Barnes G., Leclerc S., Meijer L.,
RA Kim S.H., Lockhart D.J., Schultz P.G.;
RT "Exploiting chemical libraries, structure, and genomics in the search
RT for kinase inhibitors.";
RL Science 281:533-538(1998).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
CC DURING S PHASE AND G2.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61622; CAA43807.1; -
DR EMBL; X62071; CAA43895.1; -
DR EMBL; M68520; AAA35667.1; -
DR PIR; A41227; A41227.
DR PIR; S16520; S16520.
DR PIR; S17873; S17873.
DR PDB; 1FIN; 27-JAN-97.
DR PDB; 1HCK; 07-DEC-96.
DR PDB; 1HCL; 07-DEC-96.
DR PDB; 1A01; 12-NOV-97.
DR PDB; 1JST; 11-JAN-97.
DR PDB; 1JSU; 29-JUL-97.
DR PDB; 1BUH; 09-SEP-98.
DR PDB; 1B38; 23-DEC-98.
DR PDB; 1B39; 23-DEC-98.
DR PDB; 1CKP; 13-JAN-99.
DR MM; 116953; -

```

Fri May 3 11:22:28 2002

```

CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002290; Ser_thr_kin_actsite.
CC Pfam: PF00069; pkinase; 1.
CC SMART: SM00220; S_TKC; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_DOM; 1.
CC Transferrase: Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; cell division; Mitosis; Phosphorylation; 3D-structure.
FT DOMAIN 4 286 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION.
FT MOD_RES 15 15 PHOSPHORYLATION (BY CAK).
FT MUTAGEN 14 14 T->A: INCREASE ACTIVITY 2 FOLD.
FT MUTAGEN 15 15 Y->F: INCREASE ACTIVITY 2 FOLD.
FT MUTAGEN 160 160 T->A: ABOLISHES ACTIVITY.
FT SEQUENCE 298 AA; 33929 MW; F90A0F4E70910B51 CRC64;

Query Match 34.1%; Score 567; DB 1; Length 298;
Best Local Similarity 40.0%; Pred. No. 3e-34;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGEGSYGVFKCRNKTSGOVAVKVFVESEDDPVVKKTLALREIMLKQLKH 60
DB 1 MENFQKEKIGEGTYGVYKARKNLTGTEVVALKRLDTEGVPTAIREISLLKELNH 60
QY 61 PNLVNLIEVFRKKRKHVFEYCDHTLLNELRNP-NGVADGVTKSVLWQTLQALNFCHI 119
DB 61 PNLVNLIEVFRKKRKHVFEYCDHTLLNELRNP-NGVADGVTKSVLWQTLQALNFCHI 120
QY 120 HNCIHRDKIPENILITKQGIKICDFGPAQIL-IPGDVATDYVATRWYRAPELLVGDQY 178
DB 120 HNCIHRDKIPENILITKQGIKICDFGPAQIL-IPGDVATDYVATRWYRAPELLVGDQY 180
QY 179 GSVSDIWAIGCVFAELLTGQPLWPKGSVDVQLYLIITRTGLKLIPIRHOSIFKSN 237
DB 179 GSVSDIWAIGCVFAELLTGQPLWPKGSVDVQLYLIITRTGLKLIPIRHOSIFKSN 231
QY 238 SIPE--PEDMETLEEFSDVHPV---ALNFMKGLKMPDRLTCSOLLESSYF 286
DB 238 SIPE--PEDMETLEEFSDVHPV---ALNFMKGLKMPDRLTCSOLLESSYF 286
QY 232 SMEDYKPFPPKWAQDFSKVPPLEDGDSLLSOMLHYDPNKRISAKAALAHPEF 286
DB 232 SMEDYKPFPPKWAQDFSKVPPLEDGDSLLSOMLHYDPNKRISAKAALAHPEF 286

RESULT 9
ID CC2_MAIZE STANDARD; PRT; 294 AA.
AC P23111;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG (EC 2.7.1.1.-).
GN CDC2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9119354; PubMed=2014258;
RA Colasanti J., Tyers M., Sundaresan V.;
RT *Isolation and characterization of cDNA clones encoding a functional
RL p34cdc2 homologue from Zea mays.
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY

```

```

CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002290; Ser_thr_kin_actsite.
CC Pfam: PF00069; pkinase; 1.
CC SMART: SM00220; S_TKC; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_DOM; 1.
CC Transferrase: Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; cell division; Mitosis; Phosphorylation; 3D-structure.
FT DOMAIN 4 286 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION.
FT MOD_RES 15 15 PHOSPHORYLATION (BY CAK).
FT MUTAGEN 14 14 T->A: INCREASE ACTIVITY 2 FOLD.
FT MUTAGEN 15 15 Y->F: INCREASE ACTIVITY 2 FOLD.
FT MUTAGEN 160 160 T->A: ABOLISHES ACTIVITY.
FT SEQUENCE 298 AA; 33929 MW; F90A0F4E70910B51 CRC64;

Query Match 34.1%; Score 567; DB 1; Length 298;
Best Local Similarity 40.0%; Pred. No. 3e-34;
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

QY 1 MEKYEKLAKTGEGSYGVFKCRNKTSGOVAVKVFVESEDDPVVKKTLALREIMLKQLKH 60
DB 1 MENFQKEKIGEGTYGVYKARKNLTGTEVVALKRLDTEGVPTAIREISLLKELNH 60
QY 61 PNLVNLIEVFRKKRKHVFEYCDHTLLNELRNP-NGVADGVTKSVLWQTLQALNFCHI 119
DB 61 PNLVNLIEVFRKKRKHVFEYCDHTLLNELRNP-NGVADGVTKSVLWQTLQALNFCHI 120
QY 120 HNCIHRDKIPENILITKQGIKICDFGPAQIL-IPGDVATDYVATRWYRAPELLVGDQY 178
DB 120 HNCIHRDKIPENILITKQGIKICDFGPAQIL-IPGDVATDYVATRWYRAPELLVGDQY 180
QY 179 GSVSDIWAIGCVFAELLTGQPLWPKGSVDVQLYLIITRTGLKLIPIRHOSIFKSN 237
DB 179 GSVSDIWAIGCVFAELLTGQPLWPKGSVDVQLYLIITRTGLKLIPIRHOSIFKSN 231
QY 238 SIPE--PEDMETLEEFSDVHPV---ALNFMKGLKMPDRLTCSOLLESSYF 286
DB 238 SIPE--PEDMETLEEFSDVHPV---ALNFMKGLKMPDRLTCSOLLESSYF 286
QY 232 SMEDYKPFPPKWAQDFSKVPPLEDGDSLLSOMLHYDPNKRISAKAALAHPEF 286
DB 232 SMEDYKPFPPKWAQDFSKVPPLEDGDSLLSOMLHYDPNKRISAKAALAHPEF 286

RESULT 9
ID CC2_MAIZE STANDARD; PRT; 294 AA.
AC P23111;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG (EC 2.7.1.1.-).
GN CDC2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9119354; PubMed=2014258;
RA Colasanti J., Tyers M., Sundaresan V.;
RT *Isolation and characterization of cDNA clones encoding a functional
RL p34cdc2 homologue from Zea mays.
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY

```

```

CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002290; Ser_thr_kin_actsite.
CC Pfam: PF00069; pkinase; 1.
CC SMART: SM00220; S_TKC; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_DOM; 1.
CC Transferrase: Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; cell division; Mitosis; Phosphorylation.
FT DOMAIN 4 287 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY CAK).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
FT SEQUENCE 294 AA; 33834 MW; 5063ECFCC2D3FFDD CRC64;

Query Match 34.1%; Score 566.5; DB 1; Length 294;
Best Local Similarity 37.6%; Pred. No. 3.2e-34;
Matches 117; Conservative 66; Mismatches 93; Indels 35; Gaps 9;

QY 1 MEKYEKLAKTGEGSYGVFKCRNKTSGOVAVKVFVESEDDPVVKKTLALREIMLKQLKH 60
DB 1 MEQYEKKEIGEGTYGVYKALDKATNETIALKRLKEQEDGVPTAIREISLLKEMNH 60
QY 61 PNLVNLIEVFRKKRKHVFEYCDHTLLN-----ELERNPGVADGVTKSVLWQTLQAL 114
DB 61 GNIVRLHDVHSEKRIYLVFEYLDLKKFMDSCPEFAKNPT-----LIKSYLIQILHGV 115
QY 115 NFCHTHNCIHRDKIPENILITKQ-GIHKICDFGPAQIL-IPGDVATDYVATRWYRAPELL 172
DB 116 AYCHSHRVLHRDLAPQNLIDRTNALKLADFGLARAFGIPVTRTFHEVWTLWYRAPEIL 175
QY 173 VGDYQYSSVDIWAIGCVFAELLTGQPLWPKGSVDVQLYLIITRTGLKLIPIRHOSIFKSN 232
DB 176 LGARQYSTPDVDMVSVGICFAEMVYNQKPLPGDSEIDELFKIFRILG--TPNEQS----- 227
QY 233 FPHGIS-IP-----EPEDMETLEEFSDVHPVAFNFMKGLKMPDRLTCSOLLE 282
DB 228 -WFGVSCLPDFKTAFFPRWQAQDLATV---VPLNDPAGLIDLLSKMLRYEFSKRITRAQALE 283
QY 283 SSVFDSFQBAQ 293
DB 284 HEYFKDLEVQ 294

RESULT 10
ID CDK2_CRGR STANDARD; PRT; 298 AA.
AC O55076;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.1.-).
GN CDK2.
OS Cricetulus griseus (Chinese hamster).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetulus.  
OX NCBI\_TaxID=10029;

RN	[1]	SEQUENCE FROM N.A.
RA		Ellenrieder C., Bartosch B., Lee G.Y., Murphy M., Sweeney C.,
RA		Thersgrenberg M., Hunt T., Carrington M., Jaussi R.;
RT		"The 39 kDa form of CDK2 arises through alternative splicing, is
RT		expressed in many but not all mammals, and is an active kinase.";
RL		Submitted (F8B-1998) to the EMBL/GenBank/DBJ databases.
CC	-1-	FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC		INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL
CC		DURING S PHASE AND G2.
CC	-1-	ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC		THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
CC		SIMILARITY).
CC	-1-	ALTERNATIVE PRODUCTS: 2 ISOFORMS: CDK2-ALPHA (SHOWN HERE) AND
CC		CDK2-BETA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	-1-	SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC		CDK2/CDK2 SUBSTRATE

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed, usage by for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/commercial>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR	EMRL: AJ223949; CAA11680.1; -
DR	InterPro: IPR000719; Euk_pkinase.
DR	InterPro: IPR002290; Ser_thr_kin_actsite.
DR	Pfam: PF00069; pkinase; 1.
DR	SMART: SM00220; S_TKC; 1.
DR	PROSITE: PS00107; PROTEIN_KINASE ATP; 1.
DR	PROSITE: PS00108; PROTEIN_KINASE ST; 1.
DR	PROSITE: PS0011; PROTEIN_KINASE DOM; 1.
KW	Transferase: Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KW	Cell division; Mitosis; Phosphorylation; Alternative splicing.
FT	DOMAIN 4 286 PROTEIN KINASE.
FT	NP_BIND 10 18 ATP (BY SIMILARITY).
FT	BINDING 33 33 ATP (BY SIMILARITY).
FT	ACT_SITE 127 127 BY SIMILARITY.
FT	MOD_RES 14 14 BY SIMILARITY.
FT	MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT	MOD_RES 160 160 PHOSPHORYLATION (BY SIMILARITY).
FT	MOD_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ	SEQUENCE 236 AA; 33873 MW; 3B580D8C2460CD8E CRC64;

Query Match 33.9%; Score 563; DB 1; Length 298;  
Best Local Similarity 39.7%; Pred. No. 5,8e-34;  
Matches 117; Conservative 61; Mismatches 99; Indels 18; Gaps 6;

[illegible]

```

RESULT 11
ID CC2A_ARATH STANDARD; PRT; 294 AA.
AC P24100;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG A (EC 2.7.1.-).
GN CDC2A OR CDC2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005715; PubMed=1840925;
RA Ferreira P.C.G., Hemerly A.S., Villarroel R., van Montagu M., Inze D.;
RT "The Arabidopsis functional homolog of the p34cdc2 protein kinase.";
RL plant Cell 3:531-540(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLUMBIA;
RA Hirayama T., Imajuku Y., Anai T., Matsui M., Oka A.;
RT "Identification of two cell-cycle-controlling cdc2 gene homologs in
RL Arabidopsis thaliana.";
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92316020; PubMed=1618302;
RA Imajuku Y., Hirayama T., Endoh H., Oka A.;
RT "Exon-intron organization of the Arabidopsis thaliana protein kinase
RL genes CDC2a and CDC2b.";
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339744; PubMed=1634002;
RA Inze D., Ferreira P.C.G., Hemerly A.S., van Montagu M.;
RL Biochem. Soc. Trans. 20:80-84(1992).
CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2A/CDCX SUBFAMILY

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its contents in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

```

DR      EMBL; A45397; AAB3643.1; -
DR      EMBL; M59198; AAB3281.1; -
DR      EMBL; D10850; SAA01623.1; -
DR      EMBL; X57839; CAA40971.1; -
DR      FIR; A48984; A48984.
DR      FIR; J00967; J00967.
DR      FIR; J0137; J0137.
DR      FIR; S23095; S23095.
DR      HSSP; P24941; LHCK.
DR      InterPro; IPR000719; Ser_kinase.
DR      InterPro; IPR002290; Euk_actsite.
DR      pfam; PF00069; pkinase; 1.
DR      SMART; SM00220; S_TKc; 1.

```



MBL; D17350; BAA04165.1; -  
 SSP; P24941; IAQ1.  
 nterPro: IPR000719; Euk\_pkinase.  
 nterPro: IPR002290; Ser\_thr\_kin\_actsite.  
 fam: PF00069; pkinase; 1.  
 SMART; SM00220; S\_TKC; 1.  
 ROSTE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 ROSTE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 ROSTE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
 transferase: Serine/threonine-protein kinase; ATP-binding;  
 cell cycle; Cell division; Mitosis; Phosphorylation.  
 DOMAIN 4 286 PROTEIN KINASE.  
 P\_BIND 10 18 ATP (BY SIMILARITY).  
 BINDING 33 33 ATP (BY SIMILARITY).  
 TIT SITE 127 127 BY SIMILARITY.  
 DD\_RES 14 15  
 DD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
 DD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
 DD\_RES 160 160 PHOSPHORYLATION (BY SIMILARITY).  
 DD\_RES 238 AA; 33839 MW; D1C036564C70

```

Match      33.6%; Score 559; DB 1; Length 298;
Local Similarity 39.7%; Pred No. 1.1e-33;
Cons 117; Conservative 60; Mismatches 100; Indels 18; Gaps

1 MEKYEKLAKTCEGSYGWPKCNKNTSGQVAVKFEVESEDDPVPKKIALRIMRLKOLKH 60
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
1 MENFOKEVIGEGTYGVYKAANKITGEVVALKRIDYTETGVPSTAIRSILSKELNH 60
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
61 PNINVLNLETPRRRKHHLLVFECYDHTLNELERP-NGVADGVKSIVSWQLOALNPCHI 119
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
61 PNIVKLLDVLTHTENKLYLVFELLHQDLKFMIDASAVTGIPLKSYLFOLLQGLAFCHS 120
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
20 HNCIHRDKPENILITKGQIKIKICDFGAQIL-IPGDATDYDAVTRWRAPELLVGDTQY 178
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
21 HRVLRDLKPQNLLINAECSIKLADFGLARAGVPVRTYTHEVTWLYRAPELLLGCKYY 180
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
79 GSSVDIWAIGCVFAELLTGPQLPWPKSDVDOLYLIIRTLGLKLIIPHOSIFKSNFFHGI- 237
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
31 STAVDIWSLCGIFAEMVTRRALFPDSDEIQDLFRFTRLGT-----PDEVVMPGVT 231
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
18 SIPB--PEDMWTELEKFSDDHPV-----ALNFMKGCLKNPNDRUTCSQLLESSEY 286
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
12 SMPDPKPSFPKWARODFSKVYPPLDEDGSLLSQMLHYDPNKRTISAKAALHAHPFF 286
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

RESULT	14
ID1_XENMLA	
CC ID	CC21_XENLA
AC	STANDARD; PRT; 302 AA.
DT	P35567;
DT	01-JUN-1994 (Rel. 29, Created)
DT	01-JUN-1994 (Rel. 29, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	CELL DIVISION CONTROL PROTEIN 2 HOMOLOG 1 (EC 2.7.1.-) (P34 PROTEIN KINASE).
DE	CC2X1.1.
GN	Xenopus laevis (African clawed frog).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Anphibia; Batrachia; Anura; Mesobatrachia; Pipelidae; Pipidae;
OC	Xenopodinae; Xenopus.
OX	NCBI_TaxID=8335;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Oocyte;
RC	MEDLINE=92318937; PubMed=1377775;
RA	Pickham K.M., Meyer A.N., Li J., Donoghue D.J.;
RA	"Requirement of moske protein kinase for meiotic maturation of Xenopus oocytes induced by a cdc2 mutant lacking regulatory phosphorylation sites.";
RT	Mol. Cell. Biol. 12:3192-3203(1992).
RL	[2]
RN	PHOSPHORYLATION OF THR-161.
RP	MEDLINE=93345456; PubMed=8344251;
RC	Fesquet D., Labbe J.-C., Delancourt J., Capony J.-P., Galas S., Girard F., Lorca T.T., Shuttleworth J., Doree W., Cavadore J.-C.;
RA	"The Mol5 gene encodes the catalytic subunit of a protein kinase that activates cdc2 and other cyclin-dependent kinases (CDKs) through phosphorylation of Thr161 and its homologues.";
RT	EMBO J. 12:3111-3121(1993).
RL	
CC	-1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II (BY SIMILARITY).
CC	-1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT.
CC	-1- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH A REGULATORY SUBUNIT AND WITH A CYCLIN.
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. CDC2/CDKX SUBFAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/annotation or send an email to license@isb-sib.ch).
CC	EMBL; M60680; AAA63561.1;
DR	PIR; A44349; A44349.
DR	HSP; P24941; LHCK.
DR	InterPro; IPR000719; Euk_pkinase.
DR	InterPro; IPR002290; Ser_thr_kin_actsite.
DR	Pfam; PF00069; pkinase.1.
DR	SMART; SMW0220; S-Trk; 1.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW	Transferase; Serine/threonine-protein kinase; ATP-binding;
KW	Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
FT	DOMAIN 4 287 PROTEIN KINASE.
FT	NP_BIND 10 18 ATP (BY SIMILARITY).
FT	BINDING 33 33 ATP (BY SIMILARITY).
FT	ACTION SITE 128 128 ATP (BY SIMILARITY).
FT	MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).





-----

Fri May 3 11:22:28 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:08:12 ; Search time 40.57 Seconds  
(without alignments)  
1135.711 Million cell updates/sec

Title: US-09-671-050-12  
Perfect score: 1662  
Sequence: 1 MEKYEKLAKTGEGSYGVVFK.....RKARNEGNRRROOVLPLKS 315

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues 473505  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mmc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1027.5	61.8	353	5 Q9U2H1	Q9u2h1 caenorhabdli
2	1027.5	61.8	392	5 Q9VMN3	Q9vmn3 drosophila
3	967.5	58.2	566	6 Q9TUK0	Q9ttk0 oryctolagus
4	941.5	56.6	493	4 Q92772	Q9qy12 homo sapien
5	930.5	56.0	564	11 Q9Q112	Q9qy12 mus musculus
6	930.5	56.0	568	11 Q9Q0K0	Q9quk0 mus musculus
7	928	55.8	329	11 Q9QY11	Q9qyl1 mus musculus
8	784.5	47.2	455	4 Q9P114	Q9p114 homo sapien
9	781	47.0	657	5 Q9BMG2	Q9bmg2 trypanosoma
10	769.5	46.3	1106	5 Q9GRT9	Q9grt9 leishmania
11	763.5	45.9	457	11 Q9JM02	Q9jmo2 rattus norv
12	763.5	45.9	505	11 Q9JM01	Q9jmo1 rattus norv
13	644	38.7	1104	13 Q9WER6	Q9wer6 fugu rubrip
14	641.5	38.6	997	4 Q9UJL6	Q9ujl6 homo sapien
15	579.5	34.9	294	10 Q9ZK11	Q9zrk11 tritricum ae
16	567.5	34.1	303	13 Q9DGA2	Q9dga2 oryzias jav
17	564.5	34.0	303	13 Q9DGA1	Q9dga1 oryzias jav
18	564	33.9	288	5 Q9X2D6	Q9xzd6 plasmodium
19	562	33.8	288	5 Q96820	Q96820 plasmodium

20	561.5	33.8	303	13 Q9DGD3	Q9dgd3 oryzias lat
21	560	33.7	288	5 Q96821	Q96821 plasmodium
22	559.5	33.7	303	13 Q9DGA5	Q9dga5 oryzias cur
23	559.5	33.7	303	13 Q9DGA8	Q9dga8 oryzias luz
24	558.5	33.6	294	10 P93101	P93101 chenopodium
25	556.5	33.5	294	10 Q9M307	Q9m307 arabidopsis
26	555.5	33.4	300	3 Q13379	Q13379 pneumocysti
27	555.5	33.4	300	3 Q13380	Q13380 pneumocysti
28	555.5	33.4	300	5 Q15890	Q15890 toxoplasma
29	554.5	33.4	294	10 Q82666	Q82666 brassica na
30	554.5	33.4	294	10 Q40790	Q40790 pinus conto
31	552.5	33.2	300	5 Q44000	Q44000 toxoplasma
32	551	33.2	292	13 Q9DE44	Q9de44 brachydanio
33	550.5	33.1	294	10 Q41639	Q41639 vigna aconi
34	550.5	33.1	294	10 Q43361	Q43361 picea abies
35	549.5	33.1	294	5 Q76541	Q76541 cryptospori
36	549.5	33.1	294	10 Q40484	Q40484 nicotiana t
37	549.5	33.1	294	10 Q9FUR4	Q9fur4 nicotiana t
38	549	33.0	298	5 Q27032	Q27032 theileria p
39	547.5	32.9	294	10 P93556	P93556 sesbania ro
40	547.5	32.9	300	5 Q17066	Q17066 asterina pe
41	547	32.9	298	5 Q28671	Q28671 theileria a
42	546.5	32.9	297	5 Q9TX74	Q9tx74 drosophila
43	546.5	32.9	297	5 Q9TX69	Q9tx69 drosophila
44	545.5	32.8	297	5 Q9TX73	Q9tx73 drosophila
45	544.5	32.8	294	10 Q65838	Q65838 lycopersico

## ALIGNMENTS

RESULT 1  
Q9U2H1 PRELIMINARY; PRT; 353 AA.  
ID Q9U2H1  
AC Q9U2H1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Y42A5A.4 PROTEIN.  
GN Y42A5A.4  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Steward C.A.;  
RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology.";  
RL Science 282:2012-2018(1998).  
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.  
EMBL; AL032618; CAB63367.1; -!  
HSSP; P24941; JHCL.  
DR Interpro; IPR000719; Euk\_pkinase.  
DR Interpro; IPR002290; Ser\_thr\_kin\_actsite.  
DR Pfam; PF00069; pkinase; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 353 AA; 40665 MW; 7E32260D9849CAC9 CRC64;

Query Match 61.8%; Score 1027.5; DB 5; Length 353;  
Best Local Similarity 58.8%; Pred. No. 8.9e-83;  
Matches 187; Conservative 59; Mismatches 63; Indels 9; Gaps 2;



[illegible]

```

QY 61 PNLNLEIEVFRKRKMHVFEYCDHTLLNELERNPNVADGVKRSVLSWQTLQALNFCIH 120
DB 61 ENLVNLEVCCKKKRWLVFEFVDHTILDLDLKFPLNGLDYQVQKYLFLQINGIGFCHSH 120
QY 121 NCIHRIKIPENILITKOGIIKICDFGFAQIL-IPGDATYDVATRWYRAPELLVGDTOYG 179
DB 121 NIHRDIKIPENILVSOGVVKLCDFGFAQIL-IPGDATYDVATRWYRAPELLVGDVYKG 180
QY 180 SVDIWAICVFAELLTGQPLWPKGSDVQOLYLITRTGLKLIIPRHQSIFKSNGFHGISI 239
DB 180 SVDIWAICVFAELLTGQPLWPKGSDVQOLYLITRTGLKLIIPRHQSIFKSNGFHGISI 239
QY 240 PEPEDMET--LEKFSVDHVPALNFKMGLKMPDRLTCSOLLESSYF--DSF-----Q 290
DB 240 PEPEDMET--LEKFSVDHVPALNFKMGLKMPDRLTCSOLLESSYF--DSF-----Q 290
QY 291 EQAIK--RKARN-----EGRNRRQ 309
DB 291 EQAIK--RKARN-----EGRNRRQ 309
QY 301 ELQLEKIDKARNNSLPKKSQNRKKEK 326
DB 301 ELQLEKIDKARNNSLPKKSQNRKKEK 326

RESULT 6
Q90UK0 PRELIMINARY: PRT; 568 AA.
AC Q90UK0
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SER/THR KINASE KIAMRE-DELTA (FRAGMENT).
GN Cdk12 OR KKM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ, And C57BL/6;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KIAMRE gene: variants, dual promoters, expression and
chromosomal localization.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029067; BAA88439.1; JOINED.
DR EMBL; AB029068; BAA88439.1; JOINED.
DR EMBL; AB029069; BAA88439.1; JOINED.
DR EMBL; AB029070; BAA88439.1; JOINED.
DR EMBL; AB029071; BAA88439.1; JOINED.
DR EMBL; AB029072; BAA88439.1; JOINED.
DR EMBL; AB029073; BAA88439.1; JOINED.
DR HSP; P27703; IERK.
DR MGI; MGI:1858227; Cdk12.
DR InterPro: IPR000719; Euk_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 568 AA: 64055 MW; A43B75E2E9EB86C4 CRC64;

Query Match 56.0%; Score 930.5; DB 11; Length 568;
Best Local Similarity 53.4%; Pred. No. 6.3e-74;
Matches 174; Conservative 66; Mismatches 69; Indels 17; Gaps 6;

QY 1 MEKYEKLAKTGECSGVVFKCRNKTSGOVAVKVFESDDPVVKKIALREIMLKQLKH 60
DB 1 MEKYEKLAKTGECSGVVFKCRNKTSGOVAVKVFESDDPVVKKIALREIMLKQLKH 60
QY 61 PNLNLEIEVFRKRKMHVFEYCDHTLLNELERNPNVADGVKRSVLSWQTLQALNFCIH 120
DB 61 ENLVNLEVCCKKKRWLVFEFVDHTILDLDLKFPLNGLDYQVQKYLFLQINGIGFCHSH 120

```

```

QY 121 NCIHRIKIPENILITKOGIIKICDFGFAQIL-IPGDATYDVATRWYRAPELLVGDTOYG 179
DB 121 NIHRDIKIPENILVSOGVVKLCDFGFAQIL-IPGDATYDVATRWYRAPELLVGDVYKG 180
QY 180 SVDIWAICVFAELLTGQPLWPKGSDVQOLYLITRTGLKLIIPRHQSIFKSNGFHGISI 239
DB 180 SVDIWAICVFAELLTGQPLWPKGSDVQOLYLITRTGLKLIIPRHQSIFKSNGFHGISI 239
QY 240 PEPEDMET--LEKFSVDHVPALNFKMGLKMPDRLTCSOLLESSYF--DSF-----Q 290
DB 240 PEPEDMET--LEKFSVDHVPALNFKMGLKMPDRLTCSOLLESSYF--DSF-----Q 290
QY 291 EQAIK--RKARN-----EGRNRRQ 309
DB 291 EQAIK--RKARN-----EGRNRRQ 309
QY 301 ELQLEKIDKARNNSLPKKSQNRKKEK 326
DB 301 ELQLEKIDKARNNSLPKKSQNRKKEK 326

RESULT 7
Q9QYI1 PRELIMINARY: PRT; 329 AA.
AC Q9QYI1
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SER/THR KINASE KTIAMRE-DELTA (FRAGMENT).
GN Cdk12 OR KKM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=BRIN;
RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KTIAMRE gene: variants, dual promoters, expression and
chromosomal localization.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029067; BAA88429.1; JOINED.
DR HSP; P27703; IERK.
DR MGI; MGI:1858227; Cdk12.
DR InterPro: IPR000719; Euk_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 329 AA: 38023 MW; F72EB9E65AE8585 CRC64;

Query Match 55.8%; Score 928; DB 11; Length 329;
Best Local Similarity 53.9%; Pred. No. 5e-74;
Matches 173; Conservative 63; Mismatches 73; Indels 12; Gaps 5;

QY 1 MEKYEKLAKTGECSGVVFKCRNKTSGOVAVKVFESDDPVVKKIALREIMLKQLKH 60
DB 1 MEKYEKLAKTGECSGVVFKCRNKTSGOVAVKVFESDDPVVKKIALREIMLKQLKH 60
QY 61 PNLNLEIEVFRKRKMHVFEYCDHTLLNELERNPNVADGVKRSVLSWQTLQALNFCIH 120
DB 61 ENLVNLEVCCKKKRWLVFEFVDHTILDLDLKFPLNGLDYQVQKYLFLQINGIGFCHSH 120
QY 121 NCIHRIKIPENILITKOGIIKICDFGFAQIL-IPGDATYDVATRWYRAPELLVGDTOYG 179
DB 121 NIHRDIKIPENILVSOGVVKLCDFGFAQIL-IPGDATYDVATRWYRAPELLVGDVYKG 180
QY 180 SVDIWAICVFAELLTGQPLWPKGSDVQOLYLITRTGLKLIIPRHQSIFKSNGFHGISI 239
DB 180 SVDIWAICVFAELLTGQPLWPKGSDVQOLYLITRTGLKLIIPRHQSIFKSNGFHGISI 239
QY 240 PEPEDMET--LEKFSVDHVPALNFKMGLKMPDRLTCSOLLESSYF--DSF-----Q 290
DB 240 PEPEDMET--LEKFSVDHVPALNFKMGLKMPDRLTCSOLLESSYF--DSF-----Q 290
QY 291 EQAIK--RKARN-----EGRNRRQ 309
DB 291 EQAIK--RKARN-----EGRNRRQ 309
QY 301 ELQLEKIDKARNNSLPKKSQNRKKEK 326
DB 301 ELQLEKIDKARNNSLPKKSQNRKKEK 326

```



```

Query Match          46.3%; Score 769.5; DB 5; Length 1106;
Best Local Similarity 48.3%; Pred. No. 2.7e-59;
Matches 145; Conservative 62; Mismatches 88; Indels 5; Gaps 2;

QY 1 MEKYEKLAKTGEGSYGVVFKRCNKTSQGVAVKFFVESEDDPVVKKIALREIRMLKOLKH 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MEAYETLGLGEGSYGVVFKARSRVTGKLVAKRKQTEQDEHVKTSSEVRMLQLLOH 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 PNLVNLLEVFRRKRMHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 PNVLRLDVERREGKLVLEFIDHTILQLLESTTRGFRHRELRYTYQLRGIECHNQ 120

QY 121 NCIHRLDKPENILITKOGIIKICDFGFAQILPGDAYTDYVATRWYRAPELVKDTYOG 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 NIHRDKPENILVSGSITKICDFGFARTLAAPGDVYTDYVATRWYRAPELVKLDITYG 180

QY 181 SVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIIRTLGKLIIPRHOSIFKSNGFHGISI 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 PVDVWALGCMFAELSDGQPLFGESDLQCLQMGCGVPVQRLVFIIMHNPVLYNGISFP 240

QY 241 EPDEMETLEKESDVHPVFNALFMKGCCKMNPDDRILTCQSLLSESYF--DSFQ---AQI 295
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 HTDILYTLKDRYHRESNDWIEFLSSCLHTDPAQRLTCTELMELPYFTRDGRDVEAEIR 300

RESULT 11
Q9JMW02 PRELIMINARY; PRT; 457 AA.
AC Q9JMW02;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SERINE/THREONINE KINASE NKIATRE BETA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Haq R., Randall S., Midmer M., Iafrate B., Zanke B.W.;
RT "Cloning and functional characterization of NKIATRE, a novel kinase
RT related to both MAPKs and cyclin-dependent kinases.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF112183; AAF34870.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 457 AA; 51894 MW; 99D0782794099CDA CRC64;

Query Match          45.9%; Score 763.5; DB 11; Length 457;
Best Local Similarity 46.9%; Pred. No. 2.7e-59;
Matches 145; Conservative 65; Mismatches 92; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGEGSYGVVFKRCNKTSQGVAVKFFVESEDDPVVKKIALREIRMLKOLKH 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MEAYETLGRKGEGSYGTVMCKKHKTGRIVAIIKIFYE-KPEKSVNKIATREIKFLKQFRH 59

QY 61 PNLVNLLEVFRRKRMHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ENLVNLLEVFRRKRMHLVFEYCDHTLVDELQHYCHGLESKRLRYLFLQILRAIEVLHNN 119

QY 121 NCIHRLDKPENILITKOGIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELVKDTYOG 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 NIHRDKPENILVSGSITKICDFGFARTLAAPGDVYTDYVATRWYRAPELVKLDITYG 179

QY 181 SVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIIRTLGKLIIPRHOSIFKSNGFHGISI 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 KPVDIWAIGCMITEMATGNPYPSSSDLLHKKIVLKGVNLTPHLNIFSKSPIFAGVVL 239

QY 240 PEPDEMETLEKESDVHPVFNALFMKGCCKMNPDDRILTCQSLLSESYF--DSFQ---AQI 294
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 PQVQHPKARKKYPKLNGLADIHVHACLOIDPAERISSTDLHHDYFTRDGFIEKFIPEL 299

QY 295 KKKARNEGR 303
   : : : : :
Db 300 RAKLLQEAQ 308

```

```

QY 180 SSVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIIRTLGKLIIPRHOSIFKSNGFHGISI 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 KPVDIWAIGCMITEMATGNPYPSSSDLLHKKIVLKGVNLTPHLNIFSKSPIFAGVVL 239

QY 240 PEPDEMETLEKESDVHPVFNALFMKGCCKMNPDDRILTCQSLLSESYF--DSFQ---AQI 294
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 PQVQHPKARKKYPKLNGLADIHVHACLOIDPAERISSTDLHHDYFTRDGFIEKFIPEL 299

QY 295 KKKARNEGR 303
   : : : : :
Db 300 RAKLLQEAQ 308

RESULT 12
Q9JMW01 PRELIMINARY; PRT; 505 AA.
AC Q9JMW01;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SERINE/THREONINE KINASE NKIATRE ALPHA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Haq R., Randall S., Midmer M., Iafrate B., Zanke B.W.;
RT "Cloning and functional characterization of NKIATRE, a novel kinase
RT related to both MAPKs and cyclin-dependent kinases.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF112184; AAF34871.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 505 AA; 57242 MW; 89747921FF2CE683 CRC64;

Query Match          45.9%; Score 763.5; DB 11; Length 505;
Best Local Similarity 46.9%; Pred. No. 3.1e-59;
Matches 145; Conservative 65; Mismatches 92; Indels 7; Gaps 4;

QY 1 MEKYEKLAKTGEGSYGVVFKRCNKTSQGVAVKFFVESEDDPVVKKIALREIRMLKOLKH 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MEAYETLCKVGEISYGTVMCKKHKTGRIVAIIKIFYE-KPEKSVNKIATREIKFLKQFRH 59

QY 61 PNLVNLLEVFRRKRMHLVFEYCDHTLLNELERNPNGVADGVKSVLWQTLQALNFCIH 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 60 ENLVNLLEVFRRKRMHLVFEYCDHTLVDELQHYCHGLESKRLRYLFLQILRAIEVLHNN 119

QY 121 NCIHRLDKPENILITKOGIIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELVKDTYOG 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 NIHRDKPENILVSGSITKICDFGFARTLAAPGDVYTDYVATRWYRAPELVKLDITYG 179

QY 180 SVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIIRTLGKLIIPRHOSIFKSNGFHGISI 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 KPVDIWAIGCMITEMATGNPYPSSSDLLHKKIVLKGVNLTPHLNIFSKSPIFAGVVL 239

QY 240 PEPDEMETLEKESDVHPVFNALFMKGCCKMNPDDRILTCQSLLSESYF--DSFQ---AQI 294
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 PQVQHPKARKKYPKLNGLADIHVHACLOIDPAERISSTDLHHDYFTRDGFIEKFIPEL 299

QY 295 KKKARNEGR 303
   : : : : :
Db 300 RAKLLQEAQ 308

```



us-09-671-050-12.rspt

Fri May 3 11:22:28 2002

```

RESULT 13
Q9W6R6 PRELIMINARY; PRT; 1104 AA.
ID Q9W6R6
AC Q9W6R6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SERINE-THREONINE KINASE 9.
GN STK9.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OC NCBI_TaxID=31033;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=99263230; PubMed=10330123;
RX Brunner B., Todt T., Lenzner S., Stout K., Schulz U., Ropers H.-H.,
RA Kalscheuer V.M.;
RT "Genomic structure and comparative analysis of nine Fugu genes:
RT conservation of synteny with human chromosome Xp22.2-p22.1.";
RL Genome Res. 9:437-448(1999).
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF146687; AAD28798.1; -.
DR HSP; Q16539; 1WFC.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1104 AA; 123751 MW; 9085A078B551D876 CRC64;

Query Match 38.7%; Score 644; DB 13; Length 1104;
Best Local Similarity 42.0%; Pred. No. 3.3e-48;
Matches 128; Conservative 66; Mismatches 97; Indels 14; Gaps 6;

QY 24 KTSGGVAVKFESEDDPVVKKIALREIRMLKOLKHPNLVNLIEVFRKKRKHLYFEYC 83
DB 19 KETNELERPNPGVADVTKSVLQTLQALNFCHIHNCIHRDIKPNILITKGIKIC 143
QY 84 DHTLLNELERPNPGVADVTKSVLQTLQALNFCHIHNCIHRDIKPNILITKGIKIC 143
DB 79 ERNMLELLEHPTGAPDKVRSYIYQLIKATHWCKNDIVHRDIKPNILLSSDDILKLC 138
QY 144 DFGAQLIPG-DA-YTDYVATRWYRAPELLVGDYQYSSVDIWAIGCVFAELITGQPLW 201
DB 139 DFGFARNLSEGDANDYEVATRWYRSPPELLIG-APYKAVDMWSVGCILGELSDGQPLF 197
QY 202 PGKSDVDQYLIIRTLGKLIPIHQSTFKSNGFFHGISIPEDMETLEKFSDFV-HPVAL 260
DB 198 PGSEIDQFLTQKVLGFLPAEOMKLFYNNPRFGLRFPSPVTHPQTLRRYQGLSLGML 257
QY 261 NFMKGLKMPDRLTCSQLLESSYFDSFQEAQIKRKA----RNEGRNR--QVQ 310
DB 258 DLKMLLLNPTERYLTQSLNHPATQPLQVERERAPPASPNNPRSSKRKTHIHGENT 317
QY 311 LPLAS 315
DB 318 VPTRS 322

RESULT 14
Q9JUL6 PRELIMINARY; PRT; 997 AA.
ID Q9JUL6
AC Q9JUL6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DJ245G19.2 (SERINE/THREONINE KINASE 9) (FRAGMENT).
GN STK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN SEQUENCE FROM N.A.
RP Bird C.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; Z92542; CAB38870.1; -.
DR HSP; P28523; 1A6O.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000508; Peptidase_S26.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00301; SPASE_1.1; UNKNOWN.1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 997 AA; 111939 MW; B922855AED72608D CRC64;

Query Match 38.6%; Score 641.5; DB 4; Length 997;
Best Local Similarity 43.4%; Pred. No. 4.8e-48;
Matches 124; Conservative 65; Mismatches 90; Indels 7; Gaps 4;

QY 28 QVAVKFESEDDPVVKKIALREIRMLKOLKHPNLVNLIEVFRKKRKHLYFEYCDHTL 87
DB 4 EIVAIKFKDSENEEVKTTRELKMLRTLKQENIVELKEAFRRGKLYLFEYVEKNM 63
QY 88 LNELEPNPGVADVTKSVLQTLQALNFCHIHNCIHRDIKPNILITKGIKICDFGF 147
DB 64 LELLEPNPGVPPEKVKSVIYQLIKATHWCKNDIVHRDIKPNILLSHNDVLKLCDFGF 123
QY 148 AOILIPGD--AYTDYVATRWYRAPELLVGDYQYSSVDIWAIGCVFAELITGQPLWPGKS 205
DB 124 ARNLESEGNANDYEVATRWYRSPPELLIG-APYKSDVMWSVGCILGELSDGQPLPGES 182
QY 206 DVDQLYLIIRTLGKLIPIHQSTFKSNGFFHGISIPEDMETLEKFSDFV-HPVALNEWK 264
DB 183 EIDQLFTQIKVLGFLPFAEOMKLFYNNPRFGLRFPAPVNNHPQSLRRYLGILNSVLLDLMK 242
QY 265 GCLKMNPDRLTCSQLLESSYFDSFQEAQIKRKAENGRNRNRQV 310
DB 243 NLLKLDPADRYLTQCLN---HPTQTORLLDRSPSRSAKRKPYHV 285

RESULT 15
Q9ZRI1 PRELIMINARY; PRT; 294 AA.
ID Q9ZRI1
AC Q9ZRI1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE P34CDC2.
GN CDC2TAA.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Dong C., John P.C.L.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U23409; AAD10483.1; -.
DR HSP; P24941; 1J3U.
DR Mendei; 37149; Triae;2321;37149.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.

```

DR Pfam; PF00069; pkinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 294 AA; 33996 MW; 77ABC487FA862A72 CRC64;

Query Match 34.9%; Score 579.5; DB 10; Length 294;  
Best Local Similarity 38.2%; Pred. No. 2.8e-43;  
Matches 117; Conservative 71; Mismatches 93; Indels 25; Gaps 8;

QY 1 MEKYELAKTGGSGYGVVFKCRNKTSGOVAVKKFVESEDDPVVKKIALREIRMLKOLKH 60  
Db 1 MEQYEVKIGEGTYGVVYKARDRTNETIALKKIRLEQDEGVPSTAIRISLLKEMOH 60

QY 61 PNIVNLIIEVFRKKMKHLVFEYCDHTLLNELEARNPN-GVADGVKSVLWOTLQALNFCFI 119  
Db 61 GNIVKLHDVHSEKRIWLVEYLDLDLKKFWDSCPEFAKSPALIKSYLIQLRGVAYCHS 120

QY 120 HNCIHRDIKPENILITKO-GIIKICDFGAQIL-IPGDYTDYVATRYRAPELLVGDQ 177  
Db 121 HRVLRDLKPNLLIDRTNALKLADFLARAFGIPVTFTHEVVTLWYRAPELLIGARQ 180

QY 178 YGSSVDIWAIGCVFAELLTGQPLMPGKSDVDQLYLIIRTLGLKIPRHOSIFKSNGFHGI 237  
Db 181 YSTPVDVMSVGCIFAEVNVQKPLFGDSEIDELFKIFVLG--TPNEQI-----WPGV 231

QY 238 -SIP-----EPEDMETLEEKFSVHPVAFNFMKGLKMNPDRLTCSOLLESSYFD 287  
Db 232 SSLPDYKSAPFRWQEDLATV---VPNLEPVGILDLSKMLRFPFNKRITARQALHEYFK 288

QY 288 SFQEAQ 293  
Db 289 DMEMVQ 294

Search completed: May 3, 2002, 11:08:13  
Job time: 352 sec

us-09-671-050-12.ra1

Fri May 3 11:22:27 2002

Sequence 9, Appli  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 19, Appl  
Sequence 19, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 3, Appli  
Sequence 19, Appl  
Sequence 19, Appl  
Sequence 3, Appli

1 US-08-463-090B-9  
2 US-08-874-347-18  
3 US-09-093-522-18  
4 US-08-874-347-21  
5 US-09-093-522-21  
6 US-08-318-947A-19  
7 US-08-795-303-19  
8 US-08-874-347-20  
9 US-09-093-522-20  
10 US-07-857-224B-29  
11 US-08-403-634-4  
12 US-08-306-511A-10  
13 US-08-893-274-10  
14 US-08-061-636-3  
15 US-08-874-347-19  
16 US-09-093-522-19  
17 PCT-US94-05268-3

ALIGNMENTS

RESULT 1  
US-08-874-347-25  
; Sequence 25, Application US/08874347  
; Patent No. 5863741  
; GENERAL INFORMATION:  
; APPLICANT: Linper, Andrew H.  
; APPLICANT: Leof, Edward B.  
; APPLICANT: Thomas, Charles F. P.  
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
; TITLE OF INVENTION: CARINII  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C., P.A.  
; STREET: 60 South Sixth Street, Suite 3300  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/874,347  
; FILING DATE: 13-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 07039/055001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-335-5070  
; TELEFAX: 612-288-9696  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-874-347-25

34.38; Score 570; DB 2; Length 298;  
Query Match  
Best Local Similarity 40.08; Pred. No. 5, 5e-52;  
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;

OM protein - protein search, using sw model  
Run on: May 3, 2002, 11:02:44 ; Search time 22.35 Seconds  
(without alignments)  
317.160 Million cell updates/sec

Title: us-09-671-050-12  
Perfect score: 1662  
Sequence: 1 MEKYEKLAKTGGSGYGVFEK.....RKARNEGRNRQQVLPKLS 315

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues 212252  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cqn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cqn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cqn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cqn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cqn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cqn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	570	34.3	298	2 US-08-874-347-25	Sequence 25, Appl
2	570	34.3	298	2 US-08-969-106-2	Sequence 2, Appli
3	570	34.3	298	3 US-09-093-522-25	Sequence 25, Appl
4	568.5	34.2	294	2 US-08-874-347-26	Sequence 26, Appl
5	568.5	34.2	294	3 US-09-093-522-26	Sequence 26, Appl
6	565	34.0	297	1 US-08-176-620A-16	Sequence 16, Appl
7	565	34.0	297	2 US-08-461-985-16	Sequence 16, Appl
8	557	33.5	297	2 US-08-874-347-23	Sequence 23, Appl
9	557	33.5	297	3 US-09-093-522-23	Sequence 23, Appl
10	556	33.5	297	2 US-08-874-347-24	Sequence 24, Appl
11	555.5	33.4	300	2 US-09-093-522-24	Sequence 24, Appl
12	555.5	33.4	300	2 US-08-874-347-10	Sequence 10, Appl
13	555.5	33.4	300	3 US-09-093-522-10	Sequence 10, Appl
14	555	33.2	274	2 US-07-857-224B-31	Sequence 31, Appl
15	551	33.2	274	1 US-08-318-947A-20	Sequence 20, Appl
16	551	33.2	274	2 US-08-795-303-20	Sequence 20, Appl
17	544	32.7	292	1 US-08-154-915-2	Sequence 2, Appli
18	544	32.7	292	2 US-08-464-517-38	Sequence 38, Appl
19	544	32.7	292	2 US-08-246-361A-38	Sequence 38, Appl
20	544	32.7	292	3 US-08-463-772-38	Sequence 38, Appl
21	544	32.7	292	5 PCT-US93-09945-2	Sequence 2, Appli
22	537	32.3	297	4 US-08-932-787B-21	Sequence 21, Appl
23	537	32.3	297	4 US-08-932-012C-21	Sequence 21, Appl
24	537	32.3	297	4 US-08-888-818C-21	Sequence 21, Appl
25	518.5	31.2	297	2 US-08-874-347-22	Sequence 22, Appl
26	518.5	31.2	297	3 US-09-093-522-22	Sequence 22, Appl
27	511	30.7	274	2 US-07-857-224B-30	Sequence 30, Appl

QY 1 MEKYEKLAKTEGSGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIMLKOLKH 60  
DB 1 MENFQVEKIGEGTYGVVYKARKNTLGEVVALKKIRLDTEGVPSTAIRLSLKLNLH 60  
QY 61 PNLVNLIEVFRKKRKMHLVFECDHTLLNELERNP-NGVADGVYKSVLMQTLQALNFCHEI 119  
DB 61 PNLVNLIEVFRKKRKMHLVFECDHTLLNELERNP-NGVADGVYKSVLMQTLQALNFCHEI 119  
QY 120 HNCIHRDIKPNELITKOGTIKICDFGFAQIL-IPGDAYTDYVATWYRAPELIVGDTQY 178  
DB 120 HNCIHRDIKPNELITKOGTIKICDFGFAQIL-IPGDAYTDYVATWYRAPELIVGDTQY 178  
QY 179 GSSVDIWAICGVFAELLTGQPLPCKSDVDQYLIIRLGLKIPRHQSIFKSNCFRHHI- 237  
DB 179 GSSVDIWAICGVFAELLTGQPLPCKSDVDQYLIIRLGLKIPRHQSIFKSNCFRHHI- 237  
QY 238 SIPE--PEDMETLEEKFSVDHVPV----ALNFMKGCLKMNPDRLTCSQLLESYF 286  
DB 238 SIPE--PEDMETLEEKFSVDHVPV----ALNFMKGCLKMNPDRLTCSQLLESYF 286

RESULT 2  
US-08-969-106-2  
; Sequence 2, Application US/08969106  
; Patent No. 5986055  
; GENERAL INFORMATION:  
; APPLICANT: Yang, M.  
; APPLICANT: Nandabalan, K.  
; APPLICANT: Schulz, V.  
; TITLE OF INVENTION: CDK2 INTERACTIONS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; FILING DATE: 13-NOV-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,672  
; REFERENCE/DOCKET NUMBER: 7934-057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-9741  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-969-106-2

Query Match 34.3%; Score 570; DB 2; Length 298;  
Best Local Similarity 40.0%; Pred. No. 5.5e-52;  
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;  
QY 1 MEKYEKLAKTEGSGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIMLKOLKH 60  
DB 1 MENFQVEKIGEGTYGVVYKARKNTLGEVVALKKIRLDTEGVPSTAIRLSLKLNLH 60

QY 61 PNLVNLIEVFRKKRKMHLVFECDHTLLNELERNP-NGVADGVYKSVLMQTLQALNFCHEI 119  
DB 61 PNLVNLIEVFRKKRKMHLVFECDHTLLNELERNP-NGVADGVYKSVLMQTLQALNFCHEI 119  
QY 120 HNCIHRDIKPNELITKOGTIKICDFGFAQIL-IPGDAYTDYVATWYRAPELIVGDTQY 178  
DB 120 HNCIHRDIKPNELITKOGTIKICDFGFAQIL-IPGDAYTDYVATWYRAPELIVGDTQY 178  
QY 179 GSSVDIWAICGVFAELLTGQPLPCKSDVDQYLIIRLGLKIPRHQSIFKSNCFRHHI- 237  
DB 179 GSSVDIWAICGVFAELLTGQPLPCKSDVDQYLIIRLGLKIPRHQSIFKSNCFRHHI- 237  
QY 238 SIPE--PEDMETLEEKFSVDHVPV----ALNFMKGCLKMNPDRLTCSQLLESYF 286  
DB 238 SIPE--PEDMETLEEKFSVDHVPV----ALNFMKGCLKMNPDRLTCSQLLESYF 286

RESULT 3  
US-09-093-522-25  
; Sequence 25, Application US/09093522  
; Patent No. 6015700  
; GENERAL INFORMATION:  
; APPLICANT: Limper, Andrew H.  
; APPLICANT: Leof, Edward B.  
; APPLICANT: Thomas, Charles F.  
; APPLICANT: Gustafson, Michael P.  
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C., P.A.  
; STREET: 60 South Sixth Street, Suite 3300  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; FILING DATE: 08-JUN-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/874,347  
; FILING DATE: 13-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 07039/055002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-335-5070  
; TELEFAX: 612-288-9696  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-093-522-25

Query Match 34.3%; Score 570; DB 3; Length 298;  
Best Local Similarity 40.0%; Pred. No. 5.5e-52;  
Matches 118; Conservative 60; Mismatches 99; Indels 18; Gaps 6;  
QY 1 MEKYEKLAKTEGSGYGVVFKCRNKTSGQVAVKVFVESEDDPVVKKIALREIMLKOLKH 60  
DB 1 MENFQVEKIGEGTYGVVYKARKNTLGEVVALKKIRLDTEGVPSTAIRLSLKLNLH 60

us-09-671-050-12.ra1

Fri May 3 11:22:27 2002

Db 61 GNIVRLHVDHVTSEKRIYLVFEYLDLDDLKFMWDCSPEFAKNPT-----LIKSYLYQILRGV 115  
QY 115 NFCHTHNCIHRDIPKPNILITKO-GIHKICDFGAQIL-IPGDAYTDYVATRWYRAPELL 172  
Db 116 AYCHSHRVLHRLDKPQNLIDRRNTALADGLARAFGIPVTRTFHEVVTWLTWYRAPELL 175  
QY 173 VGDYQSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIITRTGLKLIIPRHOSIFKNSG 232  
Db 176 LGSROYSTSPVDMWSVGCIFAEWNQKPLFGDSEIDELFKIFRVLG--TPNEQS----- 227  
QY 233 FFHGI-STP-----EPEDMETLEKFSDFVHPVYALNFMKGCKLKMPPDDRLTCSQLLE 282  
Db 228 -WPGVSSLPDYKSAFFKWAQDLATI---VPTPLDAGLIDLLSKMLRYEPNKRITARQALE 283  
QY 283 SSYFDSFOEAQ 293  
Db 284 HEYFKDLENVQ 294

RESULT 5  
US-09-093-522-26  
; Sequence 26, Application US/09093522  
; Patent No. 6015700  
; GENERAL INFORMATION:  
; APPLICANT: Limper, Andrew H.  
; APPLICANT: Leof, Edward B.  
; APPLICANT: Gustafson, Michael P.  
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
; TITLE OF INVENTION: CARINII  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C., P.A.  
; STREET: 60 South Sixth Street, Suite 3300  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/093,522  
; FILING DATE: 08-JUN-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/874,347  
; FILING DATE: 13-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 07039/055002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-335-5070  
; TELEFAX: 612-288-9696  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 294 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-093-522-26

Query Match 34.2%; Score 568.5; DB 3; Length 294;  
Best Local Similarity 37.6%; Pred. No. 7.8e-52;  
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 61 PNLVNLIEVFRKRKMLHVFYCDHTLLNELRNP-NGVADGVTKSVLMOTLOALNFCIH 119  
Db 61 PNLVNLIDVHTENKLYLVFEYLDLDDLKFMWDCSPEFAKNPT-----LIKSYLYQILRGV 120  
QY 120 HNCIHRDIPKPNILITKO-GIHKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDYQ 178  
Db 121 HRVLHRLDKPQNLIDRRNTALADGLARAFGIPVTRTFHEVVTWLTWYRAPELLGSKY 180  
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIITRTGLKLIIPRHOSIFKNSGFFHGI- 237  
Db 181 STAVDIWLSGCIFAEYTRRALPGDSEIDQLFRIFTIGT-----PDEVVWPGVT 231  
QY 238 SIPE--PEDMETLEKFSDFVHPV---ALNFMKGCKLKMPPDDRLTCSQLLESSTF 286  
Db 232 SMPDYKSPFKWQARDFSKVPPLEDGRLSLLSQMLHYDPNKRISAKAALAHPPF 286

RESULT 4  
US-08-874-347-26  
; Sequence 26, Application US/08874347  
; Patent No. 5863741  
; GENERAL INFORMATION:  
; APPLICANT: Limper, Andrew H.  
; APPLICANT: Leof, Edward B.  
; APPLICANT: Gustafson, Michael P.  
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
; TITLE OF INVENTION: CARINII  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C., P.A.  
; STREET: 60 South Sixth Street, Suite 3300  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/874,347  
; FILING DATE: 13-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 07039/055001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-335-5070  
; TELEFAX: 612-288-9696  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 294 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-874-347-26

Query Match 34.2%; Score 568.5; DB 2; Length 294;  
Best Local Similarity 37.6%; Pred. No. 7.8e-52;  
Matches 117; Conservative 67; Mismatches 92; Indels 35; Gaps 9;

QY 1 MEKYEKIARTGEGSYGVFKCRNKTSCQVAVKVFESDDPVVYKKTALREIRMLKOLKH 60  
Db 1 MEQYEKEKIGCTGYGVYRARDKVTNETIATKIRLQDEQGVFPSTALREISLLEKMH 60  
QY 61 PNLVNLIEVFRKRKMLHVFYCDHTLLN-----ELERNPNGVADGVTKSVLMOTLOAL 114

```
Query Match          34.0%; Score 565; DB 1; Length 297;
Best Local Similarity 39.7%; Pred. NO. 1.8e-51;
```

```

RESULT 7
US-08-461-945-16
: Sequence 16, Application US/08461985
: Patent No. 5872006
: GENERAL INFORMATION:
: APPLICANT: Boulton, Teri G.
: APPLICANT: Cobb, Melanie H.
: APPLICANT: Vancopoulos, George D.
: APPLICANT: Nye, Steven
: APPLICANT: Panayiotakis, Nikos
: TITLE OF INVENTION: A family of Map2 Protein Kinases
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/461.985
: FILING DATE: 05-JUN-1995
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/176.620
: FILING DATE: 03-JAN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Mistrock, S. Leslie
: REGISTRATION NUMBER: 18,872
: REFERENCE/DOCKET NUMBER: 6526-123
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 297 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: protein

```



```

; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-093-522-23

```

```

Query Match          33.5%; Score 557; DB 3; Length 297;
Best Local Similarity 39.1%; Pred. No. 1.3e-50;
Matches 122; Conservative 63; Mismatches 95; Indels 32; Gaps 8;

```

```

QY 1 MEKYELAKTGEGSYGVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKOLK 59
DB 1 MEDYTKIEKIGETGVVYKGRHRTTQGIIVAMKKIRLESEEG-VPSTAIRISLLKELR 59
QY 60 HPNLVNLEIFVRKRKMHVFEYCDHTLLNELEPNPG--VADGVKSVLWOTLQALNFC 117
DB 60 HPNIVSQDVLMDQSRLLIFEFLSMDLKKYLDSPPGQFMDSSLVKSYLIQLOGIVFC 119
QY 118 HPHNCIHRDIKPNILITKGIKICDFGFAQL-IPGDAYTDYVATRWYRAPELLVGD 176
DB 120 HSRVLRHDLKPNLLIDDKTKIKLADFLARAFGIPIRVYTHEVVTWRSPEVLLGSA 179
QY 177 QYGSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHOSIFKSNFGPHG 236
DB 180 RYSPVDIWSIGTIFAELATKKPLFHGDSIDQLFRIFRALG--TPNNE----- 226
QY 237 ISIPEDEMETLEEFSDVHPVA-----LNPMKGLCKMNPDDRITCSQLLESS 284
DB 227 -VWPEVESLDQYKNTFPKWKPGSLASHVKNLDENGLDLSKMLVYDPKAKRISGKMAWKHP 285
QY 285 YFDSFQEAQIKR 296
DB 286 YFDDL-DNQIKK 296

```

```

RESULT 10
US-08-874-347-24
; Sequence 24, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070

```

```

; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-874-347-24

```

```

Query Match          33.5%; Score 556; DB 2; Length 297;
Best Local Similarity 39.4%; Pred. No. 1.6e-50;
Matches 123; Conservative 61; Mismatches 96; Indels 32; Gaps 8;
QY 1 MEKYELAKTGEGSYGVFKCRNKTSGQVAVKKF-VESEDDPVVKKIALREIRMLKOLK 59
DB 1 MEDYTKIEKIGETGVVYKGRHRTTQGIIVAMKKIRLESEEG-VPSTAIRISLLKELR 59
QY 60 HPNLVNLEIFVRKRKMHVFEYCDHTLLNELEPNPG--VADGVKSVLWOTLQALNFC 117
DB 60 HPNIVSQDVLMDQSRLLIFEFLSMDLKKYLDSPPGQYMDSSLVKSYLIQLOGIVFC 119
QY 118 HPHNCIHRDIKPNILITKGIKICDFGFAQL-IPGDAYTDYVATRWYRAPELLVGD 176
DB 120 HSRVLRHDLKPNLLIDDKTKIKLADFLARAFGIPIRVYTHEVVTWRSPEVLLGSA 179
QY 177 QYGSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHOSIFKSNFGPHG 236
DB 180 RYSPVDIWSIGTIFAELATKKPLFHGDSIDQLFRIFRALG--TPNNE----- 226
QY 237 ISIPEDEMETLEEFSDVHPVA-----LNPMKGLCKMNPDDRITCSQLLESS 284
DB 227 -VWPEVESLDQYKNTFPKWKPGSLASHVKNLDENGLDLSKMLVYDPKAKRISGKMAWKHP 285
QY 285 YFDSFQEAQIKR 296
DB 286 YFDDL-DNQIKK 296

```

```

RESULT 11
US-09-093-522-24
; Sequence 24, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:

```





```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-093-522-10
```

```
Query Match 33.4%; Score 555.5; DB 3; Length 300;
Best Local Similarity 37.3%; Pred. No. 1.9e-50;
Matches 113; Conservative 72; Mismatches 91; Indels 27; Gaps 8;
```

```
QY 1 MEYKELAKTEGSGYGVFKCNKTSQGVAVKKFVESEDDPVVKKIALREIMLKQLKH 60
Db 1 MEQIQLEKIGEGYGVYKKALESTGIVALKKIRLEADEGVPSTAIRLSLLKEHN 60
QY 61 PNLVNLIEVFRKKMHLVFYCDHTLLNELERNPNVADG--VKSVMQTLQALNFC 118
Db 61 DNVRLNLIHQSRLYVFEFLDLKLYMNSIPKDMMLGAEMIKKFSQVLSGVKYCH 120
QY 119 IHNCIHRDIKPNILITKQGIKICDFGFAQIL-IPGDAYTDVYATRWYRAPELLVGDQ 177
Db 121 SHRLHRDLKPNLLIDREGNKLADFGARAFGVPGRGYTHEVVTLWYRAPEVLLGGRQ 180
QY 178 YGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLGLKLIIPRHQSIEKSNFFHGI 237
Db 181 YATALDINSIGCIFAEAMATKKPLPGDSEIDEIFRIILG--TPDENS-----WPGI 231
QY 238 -SIPE-----PEDM-ETLEEKFSVHPVALNFMKGLKMPDRLTCSQLLESYF 286
Db 232 TSYPDFKATFPKSPKNGELITELDSO---GIDLKQCLRYVPAERISAKKALDHPYF 287
QY 287 DSF 289
Db 288 DDF 290
```

```
RESULT 14
US-07-857-224B-31
; Sequence 31, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
```

```
;
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION:
; ORIGINAL SOURCE:
; ORGANISM: human
; FEATURE: Protein kinase; Table 8 Column 35
; PUBLICATION INFORMATION:
; AUTHORS:
; AUTHORS: Hanks, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988
; US-07-857-224B-31
```

```
Query Match 33.4%; Score 555; DB 2; Length 270;
Best Local Similarity 41.0%; Pred. No. 1.8e-50;
Matches 119; Conservative 58; Mismatches 87; Indels 26; Gaps 7;
```

```
QY 2 ERYEKLAETGSGYGVFKCNKTSQGVAVKKF-VESEDDPVVKKIALREIMLKQLKH 60
Db 1 EDYTKIEKIGEGYGVYKGHKTTGQVYAMKKIRLESEEG-VPSTAIRLSLLKELRH 59
QY 61 PNLVNLIEVFRKKMHLVFYCDHTLLNELERNPNG--VADGVKSVLMQTLQALNFC 118
Db 60 PNLVLODVLMDQSLRYLIFEFLSMDLKKYLDSPGQYMDSSLVKSYLQILQGVFCH 119
QY 119 IHNCIHRDIKPNILITKQGIKICDFGFAQIL-IPGDAYTDVYATRWYRAPELLVGDQ 177
Db 120 SERVLHRDLKPNLLIDDKGTIKLADFGARAFGPIRVYTHEVVTLWYRSPVLLGSSAR 179
QY 178 YGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLGLKLIIPRHQSIEKSNFFHGI 237
Db 180 YTPVDIWSIGTIFAELATKKPLFGHGDSEIDOLFRIALG--TPNNE----- 225
QY 238 SIPEFDMETLEKTSVHPVALNFMKGLKMPDRLTCSQLLESYF 287
Db 226 VMPEVESLQDLDEN-----GLDLSKMLIYDPAKRISGKMLNHPYFN 268
```

```
RESULT 15
US-08-318-947A-20
; Sequence 20, Application US/08318947A
; Patent No. 5798245
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; APPLICANT: Tian, Qingsheng
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW Suite 800
; CITY: Washington
```

us-09-671-050-12.ra

Fri May 3 11:22:27 2002

STATE: DC  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,947A  
FILING DATE: 06-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/133,530  
FILING DATE: 07-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mack, Susan J. 30,951  
REGISTRATION NUMBER: A6462  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-2920  
TELEX: 6491103  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 274 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-318-947A-20

Query Match 33.2%; Score 551; DB 1; Length 274;  
Best Local Similarity 40.6%; Pred. No. 4.9e-50;  
Matches 116; Conservative 56; Mismatches 90; Indels 24; Gaps 6;  
Qy 1 MEKYEKAKTGGEGYGVVFKCRNKTSGOVAVAKKFESEDDPVKKIALREIRMLKQLKH 60  
Db 1 MENFQKEKIGEGTYGVVYKRNKLTGEVAKKIRLDYETEGVPSTAIRSILKELNH 60  
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRNP-NGVADGVYKSVLWQTLQALNFCHI 119  
Db 61 PNIVKLLDVHTENKLYLVPELHDLKFKMDASALGTGPIPLIKSYLFQLQGLARCHS 120  
Qy 120 HNCIHRDIKPENILITKGIKICDFGAQIL-IPGDAYTDYVATRYRAPELIVGDTQY 178  
Db 121 HRVLHRLKPNLLINTEGAIKLADEGLARAFGVPVRYTHEVVTLWYRAPEILLGSKYY 180  
Qy 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTLGLKLIQRHOSIFKSNGFPHGI- 237  
Db 181 STAVKIWSLGCIFAEIMVTRALFFGDSIDQLFRIFRTLTGT-----PDEVVWFGVT 231  
Qy 238 SIPE--PEDMETLEKFSDFHVPVALNFMKGLKMMNPDRLTCSQLL 281  
Db 232 SMDYKPSFKWARQDFSKVVP-----PLDEGRSLLSQML 267

Search completed: May 3, 2002, 11:02:44  
Job time: 78 sec



Fri May 3 11:22:27 2002

Sequence 6, Appli  
Sequence 10, Appli  
Sequence 4, Appli  
Sequence 11, Appli  
Sequence 2, Appli  
Sequence 16173, A  
Sequence 13250, A  
Sequence 16190, A  
Sequence 12815, A  
Sequence 16146, A  
Sequence 509, Appl  
Sequence 731, Appl  
Sequence 4, Appli  
Sequence 10, Appli  
Sequence 2, Appli  
Sequence 8, Appli  
Sequence 209, Appl  
Sequence 1236, Appl  
Sequence 17, Appli  
Sequence 2421, Appl  
Sequence 5993, Appl  
Sequence 14, Appli  
Sequence 2, Appli  
Sequence 187, Appl  
Sequence 10259, A  
Sequence 36501, A  
Sequence 13, Appli  
Sequence 18, Appli  
Sequence 29, Appli  
Sequence 29, Appli  
Sequence 1006, Appl  
Sequence 29, Appli  
Sequence 5, Appli  
Sequence 4493, Appl  
Sequence 29536, A  
Sequence 58826, A  
Sequence 11, Appli  
Sequence 12, Appli  
Sequence 19, Appli  
Sequence 1544, Appl  
Sequence 31, Appli  
Sequence 4, Appli  
Sequence 20, Appli  
Sequence 11204, A

2	1638	98.6	324	20	US-09-671-050-6
3	1636	98.4	347	20	US-09-671-050-10
4	1612	97.0	356	20	US-09-671-050-4
5	1183.5	71.2	358	18	US-09-411-628-11
6	1140	68.6	228	18	US-09-834-496A-2
7	1027.5	61.8	392	24	US-60-167-217-16173
8	1027.5	61.8	392	24	US-60-173-464-13250
9	1027.5	61.8	392	24	US-60-191-637-16190
10	1027.5	61.8	392	24	US-60-191-681-12815
11	1027.5	61.8	392	27	US-09-614-150-16146
12	986	59.3	230	24	US-60-145-137-509
13	969	58.3	285	24	US-60-146-315-731
14	967.5	58.2	566	18	US-09-411-628-4
15	941.5	56.6	493	18	US-09-411-628-10
16	791	47.6	187	20	US-09-671-050-2
17	791	47.6	198	20	US-09-671-050-8
18	641.5	38.6	997	24	US-60-208-965-209
19	581	35.0	305	19	US-09-538-092-1236
20	581	35.0	305	28	US-09-538-092-1236
21	581	35.0	333	18	US-09-953-102A-17
22	574	34.5	352	18	US-09-488-725A-2421
23	570	34.3	298	16	US-09-488-725A-5993
24	570	34.3	298	16	US-09-266-225D-14
25	570	34.3	298	17	US-09-338-125-2
26	569	34.2	298	21	US-09-771-161A-187
27	569	34.2	224	1	PCT-US01-08631-36501
28	569	34.2	298	18	US-09-411-628-13
29	567	34.1	298	18	US-09-953-102A-18
30	567	34.1	298	18	US-09-457-040A-29
31	567	34.1	298	19	US-09-457-040B-29
32	567	34.1	298	19	US-09-538-092-1006
33	567	34.1	298	27	US-09-971-873-29
34	567	34.1	298	28	US-09-574-559A-5
35	567	34.1	298	32	PCT-US01-30732-4493
36	566.5	34.1	294	24	US-60-324-109-29536
37	566.5	34.1	294	27	US-09-708-427-58826
38	565	34.0	286	7	US-08-354-681-11
39	556	33.5	297	18	US-09-411-628-12
40	556	33.5	297	18	US-09-953-102A-19
41	556	33.5	299	27	US-09-760-446A-1544
42	555	33.4	270	3	US-07-857-224A-31
43	552.5	33.2	339	18	US-09-464-523-4
44	551	33.2	274	1	PCT-US94-11053A-20
45	546.5	32.9	297	24	US-60-167-217-11204

ALIGNMENTS

RESULT 1  
US-09-671-050-12  
; Sequence 12, Application US/09671050  
; GENERAL INFORMATION: Gregory  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: Novel Human Kinase Proteins and  
; FILE REFERENCE: Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0046-USA  
; CURRENT APPLICATION NUMBER: US/09/671,050  
; PRIOR FILING DATE: 2001-06-11  
; PRIOR APPLICATION NUMBER: US 60/156,511  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 315  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-671-050-12

OM protein - protein search, using sw model  
Run on: May 3, 2002, 11:06:49 ; Search time 205.32 Seconds  
(without alignments)  
538.891 Million cell updates/sec

Title: US-09-671-050-12  
Perfect score: 1662  
Sequence: 1 MEYKELAKTGGSYGVFEK.....RKARNEGRNRQRQVLPKLS 315  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 3516493 seqs, 351254056 residues  
Total number of hits satisfying chosen parameters: 3516493  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main: \*  
1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep.\*  
25: /cgn2\_6/ptodata/2/paa/US102\_COMB.pep.\*  
26: /cgn2\_6/ptodata/2/paa/US103\_COMB.pep.\*  
27: /cgn2\_6/ptodata/2/paa/US104\_COMB.pep.\*  
28: /cgn2\_6/ptodata/2/paa/US105\_COMB.pep.\*  
29: /cgn2\_6/ptodata/2/paa/US106\_COMB.pep.\*  
30: /cgn2\_6/ptodata/2/paa/US107\_COMB.pep.\*  
31: /cgn2\_6/ptodata/2/paa/US108\_COMB.pep.\*  
32: /cgn2\_6/ptodata/2/paa/US109\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match Length	ID	Description
1	1662	100.0	315	20 US-09-671-050-12 Sequence 12, Appl

Query Match 100.0%; Score 1662; DB 20; Length 315;  
Best Local Similarity 100.0%; Pred. No. 5.3e-142;  
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKFEVESEDDPVVKKIALREIRMLKOLKH 60  
Db 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKFEVESEDDPVVKKIALREIRMLKOLKH 60  
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNNGVADGVKSVLWQTLQALNFCCHI 120  
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNNGVADGVKSVLWQTLQALNFCCHI 120  
Qy 121 NCIHRIKPENILITKOGIIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDGTQYGS 180  
Db 121 NCIHRIKPENILITKOGIIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDGTQYGS 180  
Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHOSIFKSNFGFHGISIP 240  
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHOSIFKSNFGFHGISIP 240  
Qy 241 EPDMETLEKFSVHPVAFNMKGCLKMPDDRLTCSQLLESSYFDSFOEAIKRRKARN 300  
Db 241 EPDMETLEKFSVHPVAFNMKGCLKMPDDRLTCSQLLESSYFDSFOEAIKRRKARN 300  
Qy 301 EGRNRRRQOVLPLKS 315  
Db 301 EGRNRRRQOVLPLKS 315

## RESULT 2

US-09-671-050-6  
; Sequence 6, Application US/09671050  
; GENERAL INFORMATION:  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: Novel Human Kinase Proteins and  
; FILE REFERENCE: LEX-0046-USA  
; CURRENT APPLICATION NUMBER: US/09/671.050  
; PRIOR FILING DATE: 2001-06-11  
; PRIOR APPLICATION NUMBER: US 60/156,511  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 324  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-671-050-6

Query Match 98.6%; Score 1638; DB 20; Length 324;  
Best Local Similarity 100.0%; Pred. No. 8.3e-140;  
Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKFEVESEDDPVVKKIALREIRMLKOLKH 60  
Db 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKFEVESEDDPVVKKIALREIRMLKOLKH 60  
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNNGVADGVKSVLWQTLQALNFCCHI 120  
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNNGVADGVKSVLWQTLQALNFCCHI 120  
Qy 121 NCIHRIKPENILITKOGIIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDGTQYGS 180  
Db 121 NCIHRIKPENILITKOGIIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDGTQYGS 180

Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHOSIFKSNFGFHGISIP 240  
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHOSIFKSNFGFHGISIP 240  
Qy 241 EPDMETLEKFSVHPVAFNMKGCLKMPDDRLTCSQLLESSYFDSFOEAIKRRKARN 300  
Db 241 EPDMETLEKFSVHPVAFNMKGCLKMPDDRLTCSQLLESSYFDSFOEAIKRRKARN 300  
Qy 301 EGRNRRRQOVLPLKS 310  
Db 301 EGRNRRRQOVLPLKS 310

## RESULT 3

US-09-671-050-10  
; Sequence 10, Application US/09671050  
; GENERAL INFORMATION:  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: Novel Human Kinase Proteins and  
; FILE REFERENCE: LEX-0046-USA  
; CURRENT APPLICATION NUMBER: US/09/671.050  
; PRIOR FILING DATE: 2001-06-11  
; PRIOR APPLICATION NUMBER: US 60/156,511  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 347  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-671-050-10

Query Match 98.4%; Score 1636; DB 20; Length 347;  
Best Local Similarity 90.8%; Pred. No. 1.4e-139;  
Matches 315; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

Qy 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKFEVESEDDPVVKKIALREIRMLKOLKH 60  
Db 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGQVAVKFEVESEDDPVVKKIALREIRMLKOLKH 60  
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNNGVADGVKSVLWQTLQALNFCCHI 120  
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELRPNNGVADGVKSVLWQTLQALNFCCHI 120  
Qy 121 NCIHRIKPENILITKOGIIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDGTQYGS 180  
Db 121 NCIHRIKPENILITKOGIIKICDFGFAQILIPGDAYTDYVATRWYRAPELLVGDGTQYGS 180  
Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHOSIFKSNFGFHGISIP 240  
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLGKLIIPRHOSIFKSNFGFHGISIP 240  
Qy 218 -----GKLIIPRHOSIFKSNFGFHGISIPEDMETLEKFSVHPVAFNMKGCLK 268  
Db 241 VASQAGITGKLIIPRHOSIFKSNFGFHGISIPEDMETLEKFSVHPVAFNMKGCLK 300  
Qy 269 MNPDDLRLTCSQLLESSYFDSFOEAIKRRKARNRNRQOVLPLKS 315  
Db 301 MNPDDLRLTCSQLLESSYFDSFOEAIKRRKARNRNRQOVLPLKS 347

## RESULT 4

US-09-671-050-4  
; Sequence 4, Application US/09671050  
; GENERAL INFORMATION:  
; APPLICANT: Donoho, Gregory

; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Nehls, Michael  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: Novel Human Kinase Proteins and  
; TITLE OF INVENTION: Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0046-USA  
; CURRENT APPLICATION NUMBER: US/09/671,050  
; CURRENT FILING DATE: 2001-06-11  
; PRIOR APPLICATION NUMBER: US 60/156,511  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 356  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-671-050-4

Query Match 97.0%; Score 1612; DB 20; Length 356;  
Best Local Similarity 90.6%; Pred. No. 2.2e-137;  
Matches 310; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

Qy 1 MEKYEKLTGEGSYGVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKOLKH 60  
Db 1 MEKYEKLTGEGSYGVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKOLKH 60  
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNNGVADGVKSVLWQTLQALNFCCHI 120  
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNNGVADGVKSVLWQTLQALNFCCHI 120  
Qy 121 NCIHRIKPENILITKQGIKICDFGFAQLIPGDAYTDYVATRYRAPELLVGDQYGS 180  
Db 121 NCIHRIKPENILITKQGIKICDFGFAQLIPGDAYTDYVATRYRAPELLVGDQYGS 180  
Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTL----- 217  
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLVETGFRHVDQAGLELTSDDPPA 240  
Qy 218 -----GKLIPIHQSIKSNPFPHGISPEPEDEMETLEEKSDVHPVAFNFKMGCLK 268  
Db 241 VASQAGITGKLIPIHQSIKSNPFPHGISPEPEDEMETLEEKSDVHPVAFNFKMGCLK 300  
Qy 269 MNPDRITCSQLLESSYDFSFQEAQIKRKARNEGRRRQOV 310  
Db 301 MNPDRITCSQLLESSYDFSFQEAQIKRKARNEGRRRQOV 342

RESULT 5  
US-09-411-628-11  
; Sequence 11, Application US/09411628  
; GENERAL INFORMATION:  
; APPLICANT: University of Southern California  
; TITLE OF INVENTION: cDNA, GENOMIC, AND PREDICTED PROTEIN  
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES  
; FILE REFERENCE: 13761-707  
; CURRENT APPLICATION NUMBER: US/09/411,628  
; CURRENT FILING DATE: 1999-10-01  
; EARLIER APPLICATION NUMBER: US 60/102,906  
; EARLIER FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 358  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-411-628-11

Query Match 71.2%; Score 1183.5; DB 18; Length 358;  
Best Local Similarity 71.8%; Pred. No. 1.7e-98;

Matches 211; Conservative 41; Mismatches 41; Indels 1; Gaps 1;  
Qy 1 MEKYEKLTGEGSYGVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKOLKH 60  
Db 2 MEKYEKLTGEGSYGVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKOLKH 61  
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNNGVADGVKSVLWQTLQALNFCCHI 120  
Db 62 PNLVNLIEVFRKRKMHVFEYCDHTVHLHEDRYQGVPEHLVKSITWOTLOAVNFCCHK 121  
Qy 121 NCIHRIKPENILITKQGIKICDFGFAQLIPGDAYTDYVATRYRAPELLVGDQYGS 179  
Db 122 NCIHRIKPENILITKQGIKICDFGFAQLIPGDAYTDYVATRYRAPELLVGDQYGS 181  
Qy 180 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLGKLIPIHQSIKSNPFPHGIS 239  
Db 182 PPDVWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLGKLIPIHQSIKSNPFPHGIS 241  
Qy 240 PEPEDMETLEEKSDVHPVAFNFKMGCLKMNPDRITCSQLLESSYDFSFQEAQ 293  
Db 242 PPDVWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLGKLIPIHQSIKSNPFPHGIS 295

RESULT 6  
US-09-834-496A-2  
; Sequence 2, Application US/09834496A  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann  
; TITLE OF INVENTION: THEIR USES THEREFOR  
; FILE REFERENCE: 38152000900  
; CURRENT APPLICATION NUMBER: US/09/834,496A  
; CURRENT FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 60/196,910  
; PRIOR FILING DATE: 2000-04-13  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-496A-2

Query Match 68.6%; Score 1140; DB 28; Length 228;  
Best Local Similarity 98.2%; Pred. No. 8.4e-95;  
Matches 214; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MEKYEKLTGEGSYGVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKOLKH 60  
Db 1 MEKYEKLTGEGSYGVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKOLKH 60  
Qy 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNNGVADGVKSVLWQTLQALNFCCHI 120  
Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELEPNNGVADGVKSVLWQTLQALNFCCHI 120  
Qy 121 NCIHRIKPENILITKQGIKICDFGFAQLIPGDAYTDYVATRYRAPELLVGDQYGS 180  
Db 121 NCIHRIKPENILITKQGIKICDFGFAQLIPGDAYTDYVATRYRAPELLVGDQYGS 180  
Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLG 218  
Db 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIRTLG 218

RESULT 7  
US-60-167-217-16173  
; Sequence 16173, Application US/60167217  
; GENERAL INFORMATION:  
; APPLICANT: Li, Peter W. D.  
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES  
; THEREOF

```
; FILE REFERENCE: CL000152
; CURRENT APPLICATION NUMBER: US/60/167,217
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 23195
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16173
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Drosophila
US-60-167-217-16173
```

```
Query Match          61.8%; Score 1027.5; DB 24; Length 392;
Best Local Similarity 58.4%; Pred. No. 2.8e-84;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;

QY 1 MEKYEKAKTGEAGYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKOLKH 60
DB 1 MDRYEKLSLGEAGYGVVYKCRDRETGALVAVKRFVESEDDPAIRKIALREIRLLKLNKH 60
QY 61 PNLVNLIEVFRKRKRLHVFECYDHTLNELEARNPNGVADGVKSVLWQTLQALNFCIH 120
DB 61 PNLVLSLEVFRRKRRLHVFECYDHTLNELEARNPNGVADGVKSVLWQTLQALNFCIH 120
QY 121 NCHRDIKPENILITKOGIHKICDFGFAQILIPGDYDYVATRWYRAPELLVGDYGT 180
DB 121 GCLHRDIKPENILITKOGIHKICDFGFAQILIPGDYDYVATRWYRAPELLVGDYGT 180
QY 181 SVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIIRTLGLKIPRHOSIFKSNFGFHGISIP 240
DB 181 PVDVWAIGCLFAELVGEALWPGRSVDOLYLIIRTLGLDLPRIHQIFQNEYFKGITLP 240
QY 241 EPEDMETLEEF---SDVHPVALNEMKCKLKNPDDRLTCSQLLESSEYDSF--QEAQIK 295
DB 241 VPPTLEPLEDKMPAKSQOENPLIDFLAKCLDKDPTKRWSCERKTKHSYFDDIYAKQRELE 300
QY 296 RKARNEGRRNRQQV 310
DB 301 HVNSLEAANLKOQOL 315
```

```
RESULT 8
US-60-173-464-13250
; Sequence 13250, Application US/60/173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13250
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-13250
```

```
Query Match          61.8%; Score 1027.5; DB 24; Length 392;
Best Local Similarity 58.4%; Pred. No. 2.8e-84;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;

QY 1 MEKYEKAKTGEAGYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKOLKH 60
DB 1 MDRYEKLSLGEAGYGVVYKCRDRETGALVAVKRFVESEDDPAIRKIALREIRLLKLNKH 60
QY 61 PNLVNLIEVFRKRKRLHVFECYDHTLNELEARNPNGVADGVKSVLWQTLQALNFCIH 120
DB 61 PNLVLSLEVFRRKRRLHVFECYDHTLNELEARNPNGVADGVKSVLWQTLQALNFCIH 120
```

```
QY 121 NCHRDIKPENILITKOGIHKICDFGFAQILIPGDYDYVATRWYRAPELLVGDYGT 180
DB 121 GCLHRDIKPENILITKOGIHKICDFGFAQILIPGDYDYVATRWYRAPELLVGDYGT 180
QY 181 SVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIIRTLGLKIPRHOSIFKSNFGFHGISIP 240
DB 181 PVDVWAIGCLFAELVGEALWPGRSVDOLYLIIRTLGLDLPRIHQIFQNEYFKGITLP 240
QY 241 EPEDMETLEEF---SDVHPVALNEMKCKLKNPDDRLTCSQLLESSEYDSF--QEAQIK 295
DB 241 VPPTLEPLEDKMPAKSQOENPLIDFLAKCLDKDPTKRWSCERKTKHSYFDDIYAKQRELE 300
QY 296 RKARNEGRRNRQQV 310
DB 301 HVNSLEAANLKOQOL 315
```

```
RESULT 9
US-60-191-637-16190
; Sequence 16190, Application US/60/191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16190
; LENGTH: 392
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-16190
```

```
Query Match          61.8%; Score 1027.5; DB 24; Length 392;
Best Local Similarity 58.4%; Pred. No. 2.8e-84;
Matches 184; Conservative 65; Mismatches 61; Indels 5; Gaps 2;
```

```
QY 1 MEKYEKAKTGEAGYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKOLKH 60
DB 1 MDRYEKLSLGEAGYGVVYKCRDRETGALVAVKRFVESEDDPAIRKIALREIRLLKLNKH 60
QY 61 PNLVNLIEVFRKRKRLHVFECYDHTLNELEARNPNGVADGVKSVLWQTLQALNFCIH 120
DB 61 PNLVLSLEVFRRKRRLHVFECYDHTLNELEARNPNGVADGVKSVLWQTLQALNFCIH 120
QY 121 NCHRDIKPENILITKOGIHKICDFGFAQILIPGDYDYVATRWYRAPELLVGDYGT 180
DB 121 GCLHRDIKPENILITKOGIHKICDFGFAQILIPGDYDYVATRWYRAPELLVGDYGT 180
QY 181 SVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIIRTLGLKIPRHOSIFKSNFGFHGISIP 240
DB 181 PVDVWAIGCLFAELVGEALWPGRSVDOLYLIIRTLGLDLPRIHQIFQNEYFKGITLP 240
QY 241 EPEDMETLEEF---SDVHPVALNEMKCKLKNPDDRLTCSQLLESSEYDSF--QEAQIK 295
DB 241 VPPTLEPLEDKMPAKSQOENPLIDFLAKCLDKDPTKRWSCERKTKHSYFDDIYAKQRELE 300
QY 296 RKARNEGRRNRQQV 310
DB 301 HVNSLEAANLKOQOL 315
```

```
RESULT 10
US-60-191-681-12815
; Sequence 12815, Application US/60/191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
```







us-09-671-050-12.rapm

Fri May 3 11:22:27 2002

Db 181 KAVDWAIGCLVTEMEMGEPLFPGDSIDQLYHIMMCLGNLIPRHQELFNKNPVTAGVRL 240  
Qy 240 PEPEDMETLEKPSDYHPVALNFMKGCLKNPDDRLTCSQLLESSYF--DSF-----QEA 292  
Db 241 PEIKREPLERRYPKISEVVIDLAKKCLHIDPKRPFCAELLHHDFFQMDGFAERESQEL 300  
Qy 293 QIK--RKARN-----EGRNRRQQ 309  
Db 301 QLKVKQDARNVSLSKSONRKKKEK 324

Search completed: May 3, 2002, 11:06:50  
Job time: 324 sec



Fri May 3 11:02:09 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:10:06 ; Search time 1547.66 Seconds  
(without alignments)  
11096.471 Million cell updates/sec

Title: US-09-671-050-9  
Perfect score: 1041  
Sequence: 1 atgaaaaagtataaaaatt.....aggtaacttcgctcaaaagt 1041

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 20

Total number of hits satisfying chosen parameters: 35700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: gb\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_om.\*

20: em\_or.\*

21: em\_ov.\*

22: em\_pat.\*

23: em\_ph.\*

24: em\_pl.\*

25: em\_ro.\*

26: em\_sts.\*

27: em\_sy.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htgo\_hum.\*

31: em\_htgo\_inv.\*

32: em\_htgo\_rod.\*

33: em\_htg\_hum.\*

34: em\_htg\_inv.\*

35: em\_htg\_rod.\*

36: em\_htg\_other.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1041	100.0	1041	6	AX107720	AX107720 Sequence
2	1025	98.5	1068	6	AX107714	AX107714 Sequence
3	652	62.6	945	6	AX107722	AX107722 Sequence
4	652	62.6	972	6	AX107716	AX107716 Sequence
5	572	54.9	1819	6	AX107724	AX107724 Sequence
6	572	54.9	1819	6	AX107712	AX107712 Sequence
7	454	43.6	594	6	AX107718	AX107718 Sequence
8	454	43.6	1083	6	AX166534	AX166534 Sequence
9	368	35.4	911	6	AX056404	AX056404 Sequence
10	198	19.0	80362	9	AC079615	AC079615 Homo sapi
11	198	19.0	155028	33	AC037489	AC037489 Homo sapi
12	198	19.0	164281	2	AC092672	AC092672 Homo sapi
13	159	5.7	155028	33	AC037489	AC037489 Homo sapi
14	47	4.5	10135	9	GG0312524	AJ312524 Gorilla g
15	47	4.5	10135	9	GG0312526	AJ312526 Gorilla g
16	47	4.5	10135	9	GG0312527	AJ312527 Gorilla g
17	47	4.5	10135	9	GG0312528	AJ312528 Gorilla g
18	47	4.5	10135	9	GG0241093	AJ241093 Gorilla g
19	47	4.5	10139	9	GG0312522	AJ312522 Gorilla g
20	47	4.5	10139	9	GG0312529	AJ312529 Gorilla g
21	47	4.5	10139	9	GG0312530	AJ312530 Gorilla g
22	47	4.5	10141	9	GG0312521	AJ312521 Gorilla g
23	47	4.5	10142	9	GG0312523	AJ312523 Gorilla g
24	47	4.5	10145	9	PPA270090	AJ270090 Pan panis
25	47	4.5	10145	9	PPA270091	AJ270091 Pan panis
26	47	4.5	10145	9	PPA270093	AJ270093 Pan panis
27	47	4.5	10145	9	PPA270094	AJ270094 Pan panis
28	47	4.5	10151	9	PPA270092	AJ270092 Pan panis
29	47	4.5	10151	9	PPA270065	AJ270065 Pan trogl
30	47	4.5	10151	9	PPA270066	AJ270066 Pan trogl
31	47	4.5	10151	9	PPA270068	AJ270068 Pan trogl
32	47	4.5	10151	9	PPA270070	AJ270070 Pan trogl
33	47	4.5	10151	9	PPA270071	AJ270071 Pan trogl
34	47	4.5	10151	9	PPA270072	AJ270072 Pan trogl
35	47	4.5	10152	9	PPA241092	AJ241092 Pan trogl
36	47	4.5	10154	9	PPA270061	AJ270061 Pan trogl
37	47	4.5	10154	9	PPA270062	AJ270062 Pan trogl
38	47	4.5	10154	9	PPA270064	AJ270064 Pan trogl
39	47	4.5	10154	9	PPA270067	AJ270067 Pan trogl
40	47	4.5	10154	9	PPA270073	AJ270073 Pan trogl
41	47	4.5	10154	9	PPA270076	AJ270076 Pan trogl
42	47	4.5	10154	9	PPA270077	AJ270077 Pan trogl
43	47	4.5	10154	9	PPA270078	AJ270078 Pan trogl
44	47	4.5	10154	9	PPA270079	AJ270079 Pan trogl
45	47	4.5	10154	9	PPA270080	AJ270080 Pan trogl
46	47	4.5	10154	9	PPA270081	AJ270081 Pan trogl
47	47	4.5	10154	9	PPA270082	AJ270082 Pan trogl
48	47	4.5	10154	9	PPA270083	AJ270083 Pan trogl
49	47	4.5	10154	9	PPA270084	AJ270084 Pan trogl
50	47	4.5	10154	9	PPA270085	AJ270085 Pan trogl
51	47	4.5	10154	9	PPA270086	AJ270086 Pan trogl
52	47	4.5	10154	9	PPA270087	AJ270087 Pan trogl
53	47	4.5	10154	9	PPA270088	AJ270088 Pan trogl
54	47	4.5	10154	9	PPA270089	AJ270089 Pan trogl
55	47	4.5	10155	9	PPA270075	AJ270075 Pan trogl
56	47	4.5	10155	9	PPA270074	AJ270074 Pan trogl
57	47	4.5	61359	9	AC079396	AC079396 Homo sapi
58	47	4.5	64753	9	AC035302	AL353302 Human DNA
59	47	4.5	80668	9	AC005606	AC005606 Homo sapi
60	47	4.5	80668	9	AC005207	AC005207 Homo sapi
61	47	4.5	99411	9	AC092473	AC092473 Homo sapi
62	47	4.5	118826	2	AC092473	AC092473 Homo sapi
63	47	4.5	124821	9	AC004983	AC004983 Homo sapi
64	47	4.5	125673	9	AL356796	AL356796 Human DNA
65	47	4.5	125643	2	AC026156	AC026156 Homo sapi
66	47	4.5	127683	9	AC003683	AC003683 Homo sapi
67	47	4.5	132292	9	AL138895	AL138895 Human DNA
68	47	4.5	145673	2	AL391068	AL391068 Homo sapi
69	47	4.5	153405	2	AL357252	AL357252 Homo sapi
70	47	4.5	153452	2	AC044895	AC044895 Homo sapi

71	47	4.5 156284	2	AC036154	Homo sapi	c 144	41	3.9 174442	9	CNS057Cw	AL355865	Human chr
72	47	4.5 156294	2	AC092409	Papio cyn	c 145	41	3.9 177352	2	AL355338	AL355338	Homo sapi
73	47	4.5 160156	9	AC009362	Homo sapi	c 146	41	3.9 184737	2	AC087833	AC087833	Papio cyn
74	47	4.5 163898	2	AC025255	Homo sapi	c 147	41	3.9 188079	9	AL158209	AL158209	Homo sapi
75	47	4.5 169065	2	AF002338	Homo sapi	c 148	41	3.9 190846	9	AC006372	AC006372	Homo sapi
76	47	4.5 171627	2	AF336382	Homo sapi	c 149	41	3.9 193390	2	AC073261	AC073261	Homo sapi
77	47	4.5 171801	2	AC080162	Homo sapi	c 150	41	3.9 194265	2	AL590644	AL590644	Homo sapi
78	47	4.5 172136	2	AC023959	Homo sapi	c 151	41	3.9 197212	2	AC018723	AC018723	Homo sapi
79	47	4.5 172279	2	AC023202	Homo sapi	c 152	41	3.9 219335	9	AC005015	AC005015	Homo sapi
80	47	4.5 172609	2	AF001463	Homo sapi	c 153	41	3.9 285798	2	AL589862	AL589862	Homo sapi
81	47	4.5 172769	2	AC018628	Homo sapi	c 154	40	3.8 61615	9	AC068779	AC068779	Homo sapi
82	47	4.5 172769	2	AC018628	Homo sapi	c 155	40	3.8 73431	2	AP000575	AP000575	Homo sapi
83	47	4.5 177003	2	AC023043	Homo sapi	c 156	40	3.8 109290	2	HS838114	Y12335	Homo sapien
84	47	4.5 178828	2	AL441925	Homo sapi	c 157	40	3.8 117621	9	AC003950	AC003950	Homo sapi
85	47	4.5 182387	2	AL390246	Homo sapi	c 158	40	3.8 118218	2	AP000670	AP000670	Homo sapi
86	47	4.5 186349	2	AC073420	Homo sapi	c 159	40	3.8 128272	9	AL451125	AL451125	Human DNA
87	47	4.5 201470	2	AC073924	Homo sapi	c 160	40	3.8 149597	2	AC034271	AC034271	Homo sapi
88	47	4.5 202397	2	AC074394	Homo sapi	c 161	40	3.8 149758	2	AL591163	AL591163	Homo sapi
89	47	4.5 209855	2	AL592309	Homo sapi	c 162	40	3.8 150660	9	AL157372	AL157372	Human DNA
90	47	4.5 212074	2	AC010321	Homo sapi	c 163	40	3.8 155699	2	AL591403	AL591403	Homo sapi
91	47	4.5 216477	2	AC079882	Homo sapi	c 164	40	3.8 157442	2	AC079929	AC079929	Homo sapi
92	46	4.4 201227	2	AC092171	Homo sapi	c 165	40	3.8 157454	2	AP001462	AP001462	Homo sapi
93	44	4.2 149930	9	AL136084	Human DNA	c 166	40	3.8 162658	2	AC093081	AC093081	Homo sapi
94	44	4.2 172170	2	AC018357	Homo sapi	c 167	40	3.8 163577	9	AL590428	AL590428	Human DNA
95	44	4.2 180129	2	AC021026	Homo sapi	c 168	40	3.8 166269	2	AP001187	AP001187	Homo sapi
96	43	4.1 182834	2	AC027523	Homo sapi	c 169	40	3.8 167852	9	AC072061	AC072061	Homo sapi
97	43	4.1 194198	2	AC011774	Homo sapi	c 170	40	3.8 169466	9	CNS01BW7	AL136419	Human chr
98	43	4.1 200582	2	AC008746	Homo sapi	c 171	40	3.8 169825	9	AC012652	AC012652	Homo sapi
99	43	4.1 200582	2	AC008894	Homo sapi	c 172	40	3.8 170200	9	AC008012	AC008012	Homo sapi
100	42	4.0 49201	2	AC009100	Homo sapi	c 173	40	3.8 171980	2	AP000928	AP000928	Homo sapi
101	42	4.0 49201	2	AC009100	Homo sapi	c 174	40	3.8 178100	2	AP001558	AP001558	Homo sapi
102	42	4.0 56633	9	AL391644	Human DNA	c 175	40	3.8 179859	9	AC007563	AC007563	Homo sapi
103	42	4.0 135685	2	AC069268	Homo sapi	c 176	40	3.8 180068	2	AC074213	AC074213	Homo sapi
104	42	4.0 142973	2	AC010353	Homo sapi	c 177	40	3.8 181975	2	AC079406	AC079406	Homo sapi
105	42	4.0 145465	2	AC069399	Homo sapi	c 178	40	3.8 185820	9	AL109804	AL109804	Human DNA
106	42	4.0 175046	2	AC010531	Homo sapi	c 179	40	3.8 187483	2	AC090316	AC090316	Homo sapi
107	42	4.0 179034	33	AC009109	Homo sapi	c 180	40	3.8 187939	2	AC048338	AC048338	Homo sapi
108	42	4.0 180315	2	AC068460	Homo sapi	c 181	40	3.8 190739	9	AP001361	AP001361	Homo sapi
109	42	4.0 181609	9	AC079385	Homo sapi	c 182	40	3.8 192298	9	AP000943	AP000943	Homo sapi
110	42	4.0 188332	2	AC025521	Homo sapi	c 183	40	3.8 195334	2	AC091082	AC091082	Homo sapi
111	42	4.0 200000	2	AC006294	Homo sapi	c 184	40	3.8 200026	2	AC090368	AC090368	Homo sapi
112	41	3.9 10151	9	PT8270063	Pan trogl	c 185	40	3.8 201508	2	AC026290	AC026290	Homo sapi
113	41	3.9 16459	9	HS2497560	Homo sapi	c 186	40	3.8 203300	9	AC000134	AC000134	Homo sapi
114	41	3.9 90476	2	AF209070	Homo sapi	c 187	39	3.7 10163	9	HS241023	HS241023	Homo sapi
115	41	3.9 111857	2	HS7636F13	Homo sapi	c 188	39	3.7 10163	9	HS241024	HS241024	Homo sapi
116	41	3.9 119290	2	AC068449	Homo sapi	c 189	39	3.7 10163	9	HS241025	HS241025	Homo sapi
117	41	3.9 121417	2	AL359374	Homo sapi	c 190	39	3.7 10163	9	HS241026	HS241026	Homo sapi
118	41	3.9 121826	9	AL354935	Human DNA	c 191	39	3.7 10163	9	HS241027	HS241027	Homo sapi
119	41	3.9 130361	2	AC024911	Homo sapi	c 192	39	3.7 10163	9	HS241028	HS241028	Homo sapi
120	41	3.9 130361	2	AC024911	Homo sapi	c 193	39	3.7 10163	9	HS241029	HS241029	Homo sapi
121	41	3.9 130855	9	AC004089	Homo sapi	c 194	39	3.7 10163	9	HS241030	HS241030	Homo sapi
122	41	3.9 133814	2	AC027550	Homo sapi	c 195	39	3.7 10163	9	HS241031	HS241031	Homo sapi
123	41	3.9 140974	9	AC068533	Homo sapi	c 196	39	3.7 10163	9	HS241032	HS241032	Homo sapi
124	41	3.9 143469	2	AL161637	Homo sapi	c 197	39	3.7 10163	9	HS241033	HS241033	Homo sapi
125	41	3.9 144067	2	AL138798	Homo sapi	c 198	39	3.7 10163	9	HS241034	HS241034	Homo sapi
126	41	3.9 150747	2	AC016230	Homo sapi	c 199	39	3.7 10163	9	HS241035	HS241035	Homo sapi
127	41	3.9 150781	2	AC091961	Homo sapi	c 200	39	3.7 10163	9	HS241036	HS241036	Homo sapi
128	41	3.9 150849	2	AC020890	Homo sapi	c 201	39	3.7 10163	9	HS241037	HS241037	Homo sapi
129	41	3.9 155215	2	AL391220	Homo sapi	c 202	39	3.7 10163	9	HS241038	HS241038	Homo sapi
130	41	3.9 156604	9	AL356865	Human DNA	c 203	39	3.7 10163	9	HS241039	HS241039	Homo sapi
131	41	3.9 161109	2	AC021028	Homo sapi	c 204	39	3.7 10163	9	HS241040	HS241040	Homo sapi
132	41	3.9 161609	2	AC011823	Homo sapi	c 205	39	3.7 10163	9	HS241041	HS241041	Homo sapi
133	41	3.9 161999	2	AL133352	Homo sapi	c 206	39	3.7 10163	9	HS241042	HS241042	Homo sapi
134	41	3.9 162539	2	AL390720	Homo sapi	c 207	39	3.7 10163	9	HS241043	HS241043	Homo sapi
135	41	3.9 162609	9	AC021171	Homo sapi	c 208	39	3.7 10163	9	HS241044	HS241044	Homo sapi
136	41	3.9 164264	2	AC034197	Homo sapi	c 209	39	3.7 10163	9	HS241045	HS241045	Homo sapi
137	41	3.9 165484	2	AL590671	Homo sapi	c 210	39	3.7 10163	9	HS241046	HS241046	Homo sapi
138	41	3.9 168865	2	AL157895	Homo sapi	c 211	39	3.7 10163	9	HS241047	HS241047	Homo sapi
139	41	3.9 168865	2	AC093123	Papio cyn	c 212	39	3.7 10163	9	HS241048	HS241048	Homo sapi
140	41	3.9 168908	2	AL445383	Homo sapi	c 213	39	3.7 10163	9	HS241049	HS241049	Homo sapi
141	41	3.9 170352	2	CNS078EQ	Homo sapi	c 214	39	3.7 10163	9	HS241050	HS241050	Homo sapi
142	41	3.9 171530	9	AL358974	Homo sapi	c 215	39	3.7 10163	9	HS241051	HS241051	Homo sapi
143	41	3.9 172170	2	AC083870	Homo sapi	c 216	39	3.7 10163	9	HS241052	HS241052	Homo sapi

217	39	3.7	10163	9	HS24211053	AJ241053	Homo sapi	290	38	3.7	198486	2	AC009118	AC009118	Homo sapi
218	39	3.7	10163	9	HS24211054	AJ241054	Homo sapi	291	38	3.7	227968	9	AF053356	AF053356	Homo sapi
219	39	3.7	10163	9	HS24211055	AJ241055	Homo sapi	292	37	3.6	10531	9	AC020902	AC020902	Homo sapi
220	39	3.7	10163	9	HS24211056	AJ241056	Homo sapi	293	37	3.6	15239	2	AC090678	AC090678	Homo sapi
221	39	3.7	10163	9	HS24211057	AJ241057	Homo sapi	294	37	3.6	24232	9	AL590383	AL590383	Human DNA
222	39	3.7	10163	9	HS24211058	AJ241058	Homo sapi	295	37	3.6	33458	9	AL109656	AL109656	Human DNA
223	39	3.7	10163	9	HS24211059	AJ241059	Homo sapi	296	37	3.6	39322	9	AP001741	AP001741	Homo sapi
224	39	3.7	10163	9	HS24211060	AJ241060	Homo sapi	297	37	3.6	54666	9	AC073487	AC073487	Homo sapi
225	39	3.7	10163	9	HS24211061	AJ241061	Homo sapi	298	37	3.6	57000	9	AC069278	AC069278	Homo sapi
226	39	3.7	10163	9	HS24211062	AJ241062	Homo sapi	299	37	3.6	68688	2	AC027794	AC027794	Homo sapi
227	39	3.7	10163	9	HS24211063	AJ241063	Homo sapi	300	37	3.6	76165	2	AP000583	AP000583	Homo sapi
228	39	3.7	10163	9	HS24211064	AJ241064	Homo sapi	301	37	3.6	89626	9	AC004776	AC004776	Homo sapi
229	39	3.7	10163	9	HS24211065	AJ241065	Homo sapi	302	37	3.6	97835	9	HS292610	HS292610	Homo sapi
230	39	3.7	10163	9	HS24211066	AJ241066	Homo sapi	303	37	3.6	100000	9	AP000132	AP000132	Homo sapi
231	39	3.7	10163	9	HS24211067	AJ241067	Homo sapi	304	37	3.6	100000	9	AP000210	AP000210	Homo sapi
232	39	3.7	10163	9	HS24211068	AJ241068	Homo sapi	305	37	3.6	105412	2	AC011458	AC011458	Homo sapi
233	39	3.7	10163	9	HS24211069	AJ241069	Homo sapi	306	37	3.6	114992	9	AC008772	AC008772	Homo sapi
234	39	3.7	10163	9	HS24211070	AJ241070	Homo sapi	307	37	3.6	117564	2	HS21160H5	HS21160H5	Homo sapi
235	39	3.7	10163	9	HS24211071	AJ241071	Homo sapi	308	37	3.6	120428	2	AP000721	AP000721	Homo sapi
236	39	3.7	10163	9	HS24211072	AJ241072	Homo sapi	309	37	3.6	120578	9	AC005887	AC005887	ctdb_173_
237	39	3.7	10163	9	HS24211073	AJ241073	Homo sapi	310	37	3.6	120608	9	HS469K11	HS469K11	Homo sapi
238	39	3.7	10163	9	HS24211074	AJ241074	Homo sapi	311	37	3.6	127661	2	AP001261	AP001261	Homo sapi
239	39	3.7	10163	9	HS24211075	AJ241075	Homo sapi	312	37	3.6	130282	2	AC078973	AC078973	Homo sapi
240	39	3.7	10163	9	HS24211076	AJ241076	Homo sapi	313	37	3.6	132110	9	AP001610	AP001610	Homo sapi
241	39	3.7	10163	9	HS24211077	AJ241077	Homo sapi	314	37	3.6	133893	9	AC005005	AC005005	Homo sapi
242	39	3.7	10163	9	HS24211078	AJ241078	Homo sapi	315	37	3.6	134413	2	AC092932	AC092932	Homo sapi
243	39	3.7	10163	9	HS24211079	AJ241079	Homo sapi	316	37	3.6	138996	2	AC084024	AC084024	Homo sapi
244	39	3.7	10163	9	HS24211080	AJ241080	Homo sapi	317	37	3.6	142046	2	AC010630	AC010630	Homo sapi
245	39	3.7	10163	9	HS24211081	AJ241081	Homo sapi	318	37	3.6	143029	2	AC016319	AC016319	Homo sapi
246	39	3.7	10163	9	HS24211082	AJ241082	Homo sapi	319	37	3.6	143226	9	AL353802	AL353802	Human DNA
247	39	3.7	10163	9	HS24211083	AJ241083	Homo sapi	320	37	3.6	146017	2	AC027473	AC027473	Homo sapi
248	39	3.7	10163	9	HS24211084	AJ241084	Homo sapi	321	37	3.6	146847	9	AP000248	AP000248	Homo sapi
249	39	3.7	10163	9	HS24211085	AJ241085	Homo sapi	322	37	3.6	147172	9	AC008688	AC008688	Homo sapi
250	39	3.7	10163	9	HS24211086	AJ241086	Homo sapi	323	37	3.6	150221	9	AC004593	AC004593	Homo sapi
251	39	3.7	10163	9	HS24211087	AJ241087	Homo sapi	324	37	3.6	150893	2	AC044871	AC044871	Homo sapi
252	39	3.7	10163	9	HS24211088	AJ241088	Homo sapi	325	37	3.6	151734	2	AC021225	AC021225	Homo sapi
253	39	3.7	10163	9	HS24211089	AJ241089	Homo sapi	326	37	3.6	152448	2	AP003780	AP003780	Homo sapi
254	39	3.7	10163	9	HS24211090	AJ241090	Homo sapi	327	37	3.6	152607	2	AC024109	AC024109	Homo sapi
255	39	3.7	10163	9	HS24211091	AJ241091	Homo sapi	328	37	3.6	152679	2	AC041040	AC041040	Homo sapi
256	39	3.7	10164	9	HS24211092	AJ241092	Homo sapi	329	37	3.6	153297	2	AC023971	AC023971	Homo sapi
257	39	3.7	10164	9	HS24211093	AJ241093	Homo sapi	330	37	3.6	153788	2	AC008373	AC008373	Homo sapi
258	39	3.7	15143	9	AF107890	AJ107890	Homo sapi	331	37	3.6	155488	2	AC055727	AC055727	Homo sapi
259	39	3.7	128011	9	HS333E23	282200	Human DNA	332	37	3.6	155585	2	AL355867	AL355867	Homo sapi
260	39	3.7	163559	9	AC018635	AC018635	Homo sapi	333	37	3.6	156741	2	AC046139	AC046139	Homo sapi
261	39	3.7	166824	2	AC010655	AC010655	Homo sapi	334	37	3.6	159744	2	AP002776	AP002776	Homo sapi
262	39	3.7	173877	9	HS052112	U52112	Homo sapi	335	37	3.6	160903	2	AC026306	AC026306	Homo sapi
263	39	3.7	174424	9	HS052112	U52112	Homo sapi	336	37	3.6	161822	2	AC092954	AC092954	Homo sapi
264	39	3.7	174511	2	AP003559	AP003559	Homo sapi	337	37	3.6	162912	2	AC026058	AC026058	Homo sapi
265	39	3.7	174877	2	AP003559	AP003559	Homo sapi	338	37	3.6	163873	2	AC026080	AC026080	Homo sapi
266	39	3.7	177240	2	AF129075	AF129075	Homo sapi	339	37	3.6	166014	2	AC010223	AC010223	Homo sapi
267	39	3.7	189413	9	AF129075	AF129075	Homo sapi	340	37	3.6	167738	9	AC024884	AC024884	Homo sapi
268	39	3.7	189821	9	AL354735	AL354735	Human DNA	341	37	3.6	167738	2	AC055870	AC055870	Homo sapi
269	39	3.7	202593	2	AP003108	AP003108	Homo sapi	342	37	3.6	167914	2	AC055870	AC055870	Homo sapi
270	39	3.7	204267	9	AC020916	AC020916	Homo sapi	343	37	3.6	167914	2	AP001198	AP001198	Homo sapi
271	39	3.7	214526	9	AC020916	AC020916	Homo sapi	344	37	3.6	171386	2	AP001185	AP001185	Homo sapi
272	39	3.7	215881	2	HS21C049	AL163249	Homo sapi	345	37	3.6	171566	2	AC007736	AC007736	Homo sapi
273	39	3.7	340000	9	AL353653	AL353653	Human DNA	346	37	3.6	172476	9	AC006064	AC006064	Homo sapi
274	38	3.7	66788	9	AL353685	AL353685	Human DNA	347	37	3.6	172570	2	AL590094	AL590094	Homo sapi
275	38	3.7	84661	9	AC069281	AC069281	Homo sapi	348	37	3.6	175461	2	AP001201	AP001201	Homo sapi
276	38	3.7	141378	2	AC069281	AC069281	Homo sapi	349	37	3.6	178487	2	AC083799	AC083799	Homo sapi
277	38	3.7	153289	2	AC013533	AC013533	Homo sapi	350	37	3.6	179259	2	AL157392	AL157392	Homo sapi
278	38	3.7	155334	2	AC011156	AC011156	Homo sapi	351	37	3.6	179507	2	AC015976	AC015976	Homo sapi
279	38	3.7	156573	2	AC073075	AC073075	Homo sapi	352	37	3.6	181671	9	AC007040	AC007040	Homo sapi
280	38	3.7	159387	9	AC058784	AC058784	Homo sapi	353	37	3.6	184662	9	AC025436	AC025436	Homo sapi
281	38	3.7	163781	9	AC058784	AC058784	Homo sapi	354	37	3.6	184978	2	AL590374	AL590374	Homo sapi
282	38	3.7	164388	2	AL591647	AL591647	Homo sapi	355	37	3.6	185134	9	AC007878	AC007878	Homo sapi
283	38	3.7	168136	2	AL591647	AL591647	Homo sapi	356	37	3.6	188165	9	CNS000003	CNS000003	Human chr
284	38	3.7	168136	2	AL591647	AL591647	Homo sapi	357	37	3.6	189317	9	HSJ190J20	HSJ190J20	Human chr
285	38	3.7	173629	2	AL451077	AL451077	Homo sapi	358	37	3.6	190682	2	AC024293	AC024293	Homo sapi
286	38	3.7	186418	2	AC026898	AC026898	Homo sapi	359	37	3.6	191853	2	AC084151	AC084151	Homo sapi
287	38	3.7	190872	2	AC015799	AC015799	Homo sapi	360	37	3.6	192453	2	AC023920	AC023920	Homo sapi
288	38	3.7	193026	2	AC010158	AC010158	Homo sapi	361	37	3.6	193212	9	AC007405	AC007405	Homo sapi
289	38	3.7	194681	2	AC084057	AC084057	Homo sapi	362	37	3.6	196556	2	AC018382	AC018382	Homo sapi

C 363	37	3.6	200379	2	AC023055	AC023055 Homo sapi
C 364	37	3.6	201395	2	AC040165	AC040165 Homo sapi
C 365	37	3.6	207653	2	AL355614	AL355614 Homo sapi
C 366	37	3.6	214989	2	AC019115	AC019115 Homo sapi
C 367	37	3.6	216300	2	AC069335	AC069335 Homo sapi
C 368	37	3.6	221420	2	AC011822	AC011822 Homo sapi
C 369	37	3.6	234844	2	AL359836	AL359836 Homo sapi
C 370	37	3.6	269711	2	AP000408	AP000408 Homo sapi
C 371	37	3.6	275159	9	U82670	U82670 Homo sapien
C 372	37	3.6	340000	9	AF001711	AF001711 Homo sapi
C 373	36	3.5	6500	9	AF095743	AF095743 Homo sapi
C 374	36	3.5	7660	9	AF251125	AF251125 Homo sapi
C 375	36	3.5	10142	9	PPY312508	PPY312508 Pongo pyg
C 376	36	3.5	10142	9	PPY312516	PPY312516 Pongo pyg
C 377	36	3.5	10142	9	PPY312517	PPY312517 Pongo pyg
C 378	36	3.5	10142	9	PPY312518	PPY312518 Pongo pyg
C 379	36	3.5	10142	9	PPY312519	PPY312519 Pongo pyg
C 380	36	3.5	10144	9	PPY312514	PPY312514 Pongo pyg
C 381	36	3.5	10144	9	PPY312514	PPY312514 Pongo pyg
C 382	36	3.5	39540	9	AC000079	AC000079 Homo sapi
C 383	36	3.5	43934	9	AC000068	AC000068 Homo sapi
C 384	36	3.5	46581	9	AC008982	AC008982 Homo sapi
C 385	36	3.5	71202	2	AL603649	AL603649 Homo sapi
C 386	36	3.5	73239	2	HS3737	HS3737 Human DNA s
C 387	36	3.5	93136	2	AC022085	AC022085 Homo sapi
C 388	36	3.5	98697	9	AC004854	AC004854 Homo sapi
C 389	36	3.5	99326	2	AC007411	AC007411 Homo sapi
C 390	36	3.5	105866	2	AC021602	AC021602 Homo sapi
C 391	36	3.5	109124	2	AC011483	AC011483 Homo sapi
C 392	36	3.5	113872	9	HS5769	HS5769 Human DNA s
C 393	36	3.5	128266	9	AC004655	AC004655 Homo sapi
C 394	36	3.5	134929	9	AC007908	AC007908 Homo sapi
C 395	36	3.5	137701	2	AC073185	AC073185 Homo sapi
C 396	36	3.5	145447	2	AC012651	AC012651 Homo sapi
C 397	36	3.5	148540	9	AC004984	AC004984 Homo sapi
C 398	36	3.5	149765	9	AC079168	AC079168 Homo sapi
C 399	36	3.5	152393	2	AC079168	AC079168 Homo sapi
C 400	36	3.5	153640	9	AL445669	AL445669 Human DNA
C 401	36	3.5	153875	9	AC003682	AC003682 Homo sapi
C 402	36	3.5	154848	2	AC062992	AC062992 Homo sapi
C 403	36	3.5	155663	2	AC069238	AC069238 Homo sapi
C 404	36	3.5	157579	9	AP003493	AP003493 Human Chr
C 405	36	3.5	159693	2	AC018639	AC018639 Homo sapi
C 406	36	3.5	159854	2	AC016727	AC016727 Homo sapi
C 407	36	3.5	159854	2	AC069524	AC069524 Homo sapi
C 408	36	3.5	159873	2	AC024023	AC024023 Homo sapi
C 409	36	3.5	163320	2	AC027294	AC027294 Homo sapi
C 410	36	3.5	167395	2	AC023527	AC023527 Homo sapi
C 411	36	3.5	168843	33	AL356972	AL356972 Human DNA
C 412	36	3.5	169997	9	AC008440	AC008440 Homo sapi
C 413	36	3.5	172462	2	AL355529	AL355529 Homo sapi
C 414	36	3.5	173396	2	AL390236	AL390236 Homo sapi
C 415	36	3.5	175223	9	AC004617	AC004617 Homo sapi
C 416	36	3.5	175585	2	AC023590	AC023590 Homo sapi
C 417	36	3.5	178651	9	CNS05TDX	AL358275 Human chr
C 418	36	3.5	180195	2	AC027089	AC027089 Homo sapi
C 419	36	3.5	182776	9	AC016748	AC016748 Homo sapi
C 420	36	3.5	183600	2	AC093131	AC093131 Papio cyn
C 421	36	3.5	184414	2	AC092135	AC092135 Homo sapi
C 422	36	3.5	188934	2	AC072040	AC072040 Homo sapi
C 423	36	3.5	191422	2	AC022318	AC022318 Homo sapi
C 424	36	3.5	191422	2	AC022017	AC022017 Homo sapi
C 425	36	3.5	195969	2	AC017107	AC017107 Homo sapi
C 426	36	3.5	195969	2	AC067730	AC067730 Homo sapi
C 427	36	3.5	195014	2	AC073849	AC073849 Homo sapi
C 428	36	3.5	195014	2	AL359764	AL359764 Homo sapi
C 429	36	3.5	201123	2	AL139410	AL139410 Homo sapi
C 430	36	3.5	202200	2	AC008744	AC008744 Homo sapi
C 431	36	3.5	216935	2	AC083813	AC083813 Homo sapi
C 432	36	3.5	217346	2	AC027602	AC027602 Homo sapi
C 433	36	3.5	217346	2	AC011005	AC011005 Homo sapi
C 434	36	3.5	217615	9	AC011605	AC011605 Homo sapi
C 435	36	3.5	217615	9	AC011605	AC011605 Homo sapi
C 436	36	3.5	219820	9	AC008738	AC008738 Homo sapi
C 437	36	3.5	230000	9	AF243527	AF243527 Homo sapi
C 438	36	3.4	10141	9	PPX270095	PPX270095 Pongo pyg
C 439	35	3.4	10142	9	PPX312509	PPX312509 Pongo pyg
C 440	35	3.4	10143	9	PPX312511	PPX312511 Pongo pyg
C 441	35	3.4	10143	9	PPX312515	PPX312515 Pongo pyg
C 442	35	3.4	10144	9	PPX312513	PPX312513 Pongo pyg
C 443	35	3.4	10144	9	PPX312520	PPX312520 Pongo pyg
C 444	35	3.4	36339	2	HS83387	HS83387 Human DNA s
C 445	35	3.4	47139	2	AC090452	AC090452 Homo sapi
C 446	35	3.4	73560	9	HS066083	HS066083 Human conti
C 447	35	3.4	74853	9	AC007900	AC007900 Homo sapi
C 448	35	3.4	76241	9	HS033644	HS033644 Human DNA
C 449	35	3.4	85116	9	HS032488	HS032488 Human DNA
C 450	35	3.4	92510	9	HS39083	HS39083 Human DNA s
C 451	35	3.4	103863	9	AC006596	AC006596 Homo sapi
C 452	35	3.4	111183	9	AC008950	AC008950 Homo sapi
C 453	35	3.4	112314	9	AC004002	AC004002 Human BAC
C 454	35	3.4	113297	9	AL138762	AL138762 Human DNA
C 455	35	3.4	114026	9	AL1365201	AL1365201 Human DNA
C 456	35	3.4	122839	9	AL161898	AL161898 Human DNA
C 457	35	3.4	128880	2	AL136458	AL136458 Homo sapi
C 458	35	3.4	129098	9	HS46H23	HS46H23 Human DNA s
C 459	35	3.4	129158	2	AC090170	AC090170 Homo sapi
C 460	35	3.4	135734	2	AC002346	AC002346 Homo sapi
C 461	35	3.4	138538	9	AC011445	AC011445 Homo sapi
C 462	35	3.4	139015	2	AC068016	AC068016 Homo sapi
C 463	35	3.4	145947	9	AL353588	AL353588 Human DNA s
C 464	35	3.4	147708	9	HS179M20	HS179M20 Human DNA
C 465	35	3.4	148271	9	AL392046	AL392046 Homo sapi
C 466	35	3.4	150435	2	AC069579	AC069579 Homo sapi
C 467	35	3.4	150435	2	AC068373	AC068373 Homo sapi
C 468	35	3.4	155186	2	AL355994	AL355994 Homo sapi
C 469	35	3.4	157095	2	AL355994	AL355994 Homo sapi
C 470	35	3.4	157365	9	AC025754	AC025754 Homo sapi
C 471	35	3.4	160725	2	HS019B11	HS019B11 Human DNA
C 472	35	3.4	162486	2	AC044779	AC044779 Homo sapi
C 473	35	3.4	164034	9	AC011489	AC011489 Homo sapi
C 474	35	3.4	169652	2	AC090668	AC090668 Homo sapi
C 475	35	3.4	170790	2	AC025609	AC025609 Homo sapi
C 476	35	3.4	171206	2	AC026573	AC026573 Homo sapi
C 477	35	3.4	172090	2	AC018477	AC018477 Homo sapi
C 478	35	3.4	178574	2	AL591026	AL591026 Homo sapi
C 479	35	3.4	178953	2	AC026626	AC026626 Homo sapi
C 480	35	3.4	179667	2	AC025681	AC025681 Homo sapi
C 481	35	3.4	180615	9	AC053543	AC053543 Homo sapi
C 482	35	3.4	186261	2	AC016602	AC016602 Homo sapi
C 483	35	3.4	186686	2	AL162739	AL162739 Homo sapi
C 484	35	3.4	189010	2	AC092835	AC092835 Homo sapi
C 485	35	3.4	193323	2	AC073842	AC073842 Homo sapi
C 486	35	3.4	193514	9	AC010530	AC010530 Homo sapi
C 487	35	3.4	197326	2	AC022966	AC022966 Homo sapi
C 488	35	3.4	199019	2	AL359764	AL359764 Homo sapi
C 489	35	3.4	201883	2	AC009972	AC009972 Homo sapi
C 490	35	3.4	203650	2	AC092296	AC092296 Homo sapi
C 491	35	3.4	208337	2	AC068400	AC068400 Homo sapi
C 492	34	3.3	210672	2	HS179115	HS179115 Human sapien
C 493	34	3.3	26459	9	HSBA775A3	HSBA775A3 Human DNA
C 494	34	3.3	28505	9	AL136301	AL136301 Human DNA
C 495	34	3.3	69213	9	AC010191	AC010191 Homo sapi
C 496	34	3.3	82412	2	AC040940	AC040940 Homo sapi
C 497	34	3.3	94742	9	AC006263	AC006263 Homo sapi
C 498	34	3.3	95953	9	HS3857M17	HS3857M17 Human DNA
C 499	34	3.3	110000	2	AL355364	AL355364 Continuation (2 of
C 500	34	3.3	140092	9	AL162615	AL162615 Human DNA
C 501	34	3.3	145966	9	AC006451	AC006451 Homo sapi
C 502	34	3.3	146028	9	AC021573	AC021573 Homo sapi
C 503	34	3.3	152283	9	AC024083	AC024083 Homo sapi
C 504	34	3.3	156005	2	AL158136	AL158136 Homo sapi
C 505	34	3.3	157633	9	AC008687	AC008687 Homo sapi
C 506	34	3.3	158297	2	AL157820	AL157820 Homo sapi
C 507	34	3.3	158594	2	AC087768	AC087768 Homo sapi
C 508	34	3.3	159572	33	AC022932	AC022932 Homo sapi



us-09-671-050-9.oli20.rge

Fri May 3 11:02:09 2002

c 509	34	3.3	159983	2	AC013713	Homo sapi	AC013713 Homo sapi
c 510	34	3.3	162847	2	AC025980	Homo sapi	AC025980 Homo sapi
c 511	34	3.3	164917	2	AL592064	Homo sapi	AL592064 Homo sapi
c 512	34	3.3	165611	2	AC011694	Homo sapi	AC011694 Homo sapi
c 513	34	3.3	166354	2	AC078984	Homo sapi	AC078984 Homo sapi
c 514	34	3.3	167563	9	AC069280	Homo sapi	AC069280 Homo sapi
c 515	34	3.3	168247	9	AC005014	Homo sapi	AC005014 Homo sapi
c 516	34	3.3	171007	2	AC090421	Homo sapi	AC090421 Homo sapi
c 517	34	3.3	171176	2	AC060796	Homo sapi	AC060796 Homo sapi
c 518	34	3.3	171338	2	AL157875	Homo sapi	AL157875 Homo sapi
c 519	34	3.3	171998	2	AC092648	Homo sapi	AC092648 Homo sapi
c 520	34	3.3	174761	2	AL391650	Homo sapi	AL391650 Homo sapi
c 521	34	3.3	176330	2	AC016219	Homo sapi	AC016219 Homo sapi
c 522	34	3.3	178630	2	AC011611	Homo sapi	AC011611 Homo sapi
c 523	34	3.3	179085	2	AC084291	Homo sapi	AC084291 Homo sapi
c 524	34	3.3	179260	9	AC015987	Homo sapi	AC015987 Homo sapi
c 525	34	3.3	181891	9	AC009154	Homo sapi	AC009154 Homo sapi
c 526	34	3.3	181914	9	AC016916	Homo sapi	AC016916 Homo sapi
c 527	34	3.3	183451	9	AC005972	Homo sapi	AC005972 Homo sapi
c 528	34	3.3	185380	2	AC016863	Homo sapi	AC016863 Homo sapi
c 529	34	3.3	189938	2	AC023601	Homo sapi	AC023601 Homo sapi
c 530	34	3.3	192254	2	AC010741	Homo sapi	AC010741 Homo sapi
c 531	34	3.3	192760	2	AL355586	Homo sapi	AL355586 Homo sapi
c 532	34	3.3	195439	2	AL390774	Homo sapi	AL390774 Homo sapi
c 533	34	3.3	197055	2	AC027705	Homo sapi	AC027705 Homo sapi
c 534	34	3.3	199515	2	AC018475	Homo sapi	AC018475 Homo sapi
c 535	34	3.3	199752	2	AL589942	Homo sapi	AL589942 Homo sapi
c 536	34	3.3	210096	2	AL589942	Homo sapi	AL589942 Homo sapi
c 537	34	3.3	214653	2	AL161790	Homo sapi	AL161790 Homo sapi
c 538	34	3.3	215584	2	AL390779	Homo sapi	AL390779 Homo sapi
c 539	34	3.3	2321260	2	AL160172	Homo sapi	AL160172 Homo sapi
c 540	34	3.3	271144	9	HSXDPB	Homo sapi	HSXDPB Homo sapi
c 541	34	3.3	312283	9	AF205588	Homo sapi	AF205588 Homo sapi
c 542	33	3.2	752	9	BC008313	Homo sapi	BC008313 Homo sapi
c 543	33	3.2	5507	9	AB011542	Homo sapi	AB011542 Homo sapi
c 544	33	3.2	32198	9	AF229163	Homo sapi	AF229163 Homo sapi
c 545	33	3.2	36800	9	HS313D11	Homo sapi	HS313D11 Homo sapi
c 546	33	3.2	49386	2	AC021257	Homo sapi	AC021257 Homo sapi
c 547	33	3.2	73029	9	HSBA3259J7	Homo sapi	HSBA3259J7 Homo sapi
c 548	33	3.2	76200	9	AC008616	Homo sapi	AC008616 Homo sapi
c 549	33	3.2	80419	9	AL139826	Homo sapi	AL139826 Homo sapi
c 550	33	3.2	83798	9	HS1013A22	Homo sapi	HS1013A22 Homo sapi
c 551	33	3.2	86453	9	HS1049G16	Homo sapi	HS1049G16 Homo sapi
c 552	33	3.2	87375	9	AF000289	Homo sapi	AF000289 Homo sapi
c 553	33	3.2	89016	9	HS86C11	Homo sapi	HS86C11 Homo sapi
c 554	33	3.2	90248	2	AF003326	Homo sapi	AF003326 Homo sapi
c 555	33	3.2	92797	9	AF038458	Homo sapi	AF038458 Homo sapi
c 556	33	3.2	94960	9	AC000066	Homo sapi	AC000066 Homo sapi
c 557	33	3.2	100000	9	AP000042	Homo sapi	AP000042 Homo sapi
c 558	33	3.2	100000	9	AP000049	Homo sapi	AP000049 Homo sapi
c 559	33	3.2	100000	9	AP000110	Homo sapi	AP000110 Homo sapi
c 560	33	3.2	100000	9	AP000116	Homo sapi	AP000116 Homo sapi
c 561	33	3.2	100000	9	AP000186	Homo sapi	AP000186 Homo sapi
c 562	33	3.2	100000	9	AP000192	Homo sapi	AP000192 Homo sapi
c 563	33	3.2	100000	9	AP000503	Homo sapi	AP000503 Homo sapi
c 564	33	3.2	102151	9	HS979N1	Homo sapi	HS979N1 Homo sapi
c 565	33	3.2	108523	9	AC000159	Homo sapi	AC000159 Homo sapi
c 566	33	3.2	109867	9	AC004898	Homo sapi	AC004898 Homo sapi
c 567	33	3.2	110633	9	AL353710	Homo sapi	AL353710 Homo sapi
c 568	33	3.2	111035	9	AC022405	Homo sapi	AC022405 Homo sapi
c 569	33	3.2	111990	9	AL360294	Homo sapi	AL360294 Homo sapi
c 570	33	3.2	112309	9	AC003025	Homo sapi	AC003025 Homo sapi
c 571	33	3.2	118030	2	AP000667	Homo sapi	AP000667 Homo sapi
c 572	33	3.2	121442	2	AL137860	Homo sapi	AL137860 Homo sapi
c 573	33	3.2	12530	2	AL356369	Homo sapi	AL356369 Homo sapi
c 574	33	3.2	125242	9	AP000311	Homo sapi	AP000311 Homo sapi
c 575	33	3.2	125291	9	AP000646	Homo sapi	AP000646 Homo sapi
c 576	33	3.2	125350	9	AC020768	Homo sapi	AC020768 Homo sapi
c 577	33	3.2	128230	9	HS313L4	Homo sapi	HS313L4 Homo sapi
c 578	33	3.2	128559	9	AC004677	Homo sapi	AC004677 Homo sapi
c 579	33	3.2	128899	9	AL391139	Homo sapi	AL391139 Homo sapi
c 580	33	3.2	129502	9	HSXJ726C3	Homo sapi	HSXJ726C3 Homo sapi
c 581	33	3.2	130467	9	HS798A17	Homo sapi	HS798A17 Homo sapi
c 582	33	3.2	131314	2	AP000817	Homo sapi	AP000817 Homo sapi
c 583	33	3.2	133683	9	AC084857	Homo sapi	AC084857 Homo sapi
c 584	33	3.2	134515	9	HS354J5	Homo sapi	HS354J5 Homo sapi
c 585	33	3.2	134601	2	AL450106	Homo sapi	AL450106 Homo sapi
c 586	33	3.2	136572	9	AL113384	Homo sapi	AL113384 Homo sapi
c 587	33	3.2	137833	9	CNS01DV6	Homo sapi	CNS01DV6 Homo sapi
c 588	33	3.2	138094	2	AC009417	Homo sapi	AC009417 Homo sapi
c 589	33	3.2	138627	2	AP000591	Homo sapi	AP000591 Homo sapi
c 590	33	3.2	140606	2	AL353761	Homo sapi	AL353761 Homo sapi
c 591	33	3.2	141989	2	AC089439	Homo sapi	AC089439 Homo sapi
c 592	33	3.2	142092	9	AF139813	Homo sapi	AF139813 Homo sapi
c 593	33	3.2	143769	9	AC004914	Homo sapi	AC004914 Homo sapi
c 594	33	3.2	143813	9	AC002994	Homo sapi	AC002994 Homo sapi
c 595	33	3.2	144200	9	AC022638	Homo sapi	AC022638 Homo sapi
c 596	33	3.2	144676	9	HS390013	Homo sapi	HS390013 Homo sapi
c 597	33	3.2	145414	9	HS392M18	Homo sapi	HS392M18 Homo sapi
c 598	33	3.2	147094	2	AC012027	Homo sapi	AC012027 Homo sapi
c 599	33	3.2	147925	2	AC013320	Homo sapi	AC013320 Homo sapi
c 600	33	3.2	149733	2	AC087568	Homo sapi	AC087568 Homo sapi
c 601	33	3.2	149765	9	AC004984	Homo sapi	AC004984 Homo sapi
c 602	33	3.2	151310	2	AC091980	Homo sapi	AC091980 Homo sapi
c 603	33	3.2	153284	2	AL136233	Homo sapi	AL136233 Homo sapi
c 604	33	3.2	156657	2	AL136233	Homo sapi	AL136233 Homo sapi
c 605	33	3.2	157662	9	AC023537	Homo sapi	AC023537 Homo sapi
c 606	33	3.2	158067	2	AC024423	Homo sapi	AC024423 Homo sapi
c 607	33	3.2	158287	2	AP002961	Homo sapi	AP002961 Homo sapi
c 608	33	3.2	158475	2	CNS01DXB	Homo sapi	CNS01DXB Homo sapi
c 609	33	3.2	158802	2	AC015589	Homo sapi	AC015589 Homo sapi
c 610	33	3.2	158802	33	AC015589	Homo sapi	AC015589 Homo sapi
c 611	33	3.2	158802	33	AC015589	Homo sapi	AC015589 Homo sapi
c 612	33	3.2	158802	33	AC015589	Homo sapi	AC015589 Homo sapi
c 613	33	3.2	158802	33	AC015589	Homo sapi	AC015589 Homo sapi
c 614	33	3.2	159624	2	AC011021	Homo sapi	AC011021 Homo sapi
c 615	33	3.2	159635	2	AC011009	Homo sapi	AC011009 Homo sapi
c 616	33	3.2	160156	2	AL591852	Homo sapi	AL591852 Homo sapi
c 617	33	3.2	161031	9	AC079610	Homo sapi	AC079610 Homo sapi
c 618	33	3.2	161548	9	AP001929	Homo sapi	AP001929 Homo sapi
c 619	33	3.2	162453	2	AC087778	Homo sapi	AC087778 Homo sapi
c 620	33	3.2	162589	2	AC015717	Homo sapi	AC015717 Homo sapi
c 621	33	3.2	162984	2	AL590642	Homo sapi	AL590642 Homo sapi
c 622	33	3.2	164426	9	AL355495	Homo sapi	AL355495 Homo sapi
c 623	33	3.2	166118	9	AC005225	Homo sapi	AC005225 Homo sapi
c 624	33	3.2	166774	9	AL161787	Homo sapi	AL161787 Homo sapi
c 625	33	3.2	166894	2	AL161787	Homo sapi	AL161787 Homo sapi
c 626	33	3.2	168119	9	AC018719	Homo sapi	AC018719 Homo sapi
c 627	33	3.2	168244	9	AL133344	Homo sapi	AL133344 Homo sapi
c 628	33	3.2	169223	2	AL353675	Homo sapi	AL353675 Homo sapi
c 629	33	3.2	169812	9	AL391315	Homo sapi	AL391315 Homo sapi
c 630	33	3.2	169863	2	AL591467	Homo sapi	AL591467 Homo sapi
c 631	33	3.2	169984	2	AL161789	Homo sapi	AL161789 Homo sapi
c 632	33	3.2	170008	9	HSN010770	Homo sapi	HSN010770 Homo sapi
c 633	33	3.2	170281	9	AC026410	Homo sapi	AC026410 Homo sapi
c 634	33	3.2	170591	9	AC026410	Homo sapi	AC026410 Homo sapi
c 635	33	3.2	171034	2	AP001325	Homo sapi	AP001325 Homo sapi
c 636	33	3.2	171759	2	AC073650	Homo sapi	AC073650 Homo sapi
c 637	33	3.2	171980	2	AP000928	Homo sapi	AP000928 Homo sapi
c 638	33	3.2	172105	2	AC021197	Homo sapi	AC021197 Homo sapi
c 639	33	3.2	172138	2	AC064796	Homo sapi	AC064796 Homo sapi
c 640	33	3.2	172585	9	AC090527	Homo sapi	AC090527 Homo sapi
c 641	33	3.2	174733	2	AC016386	Homo sapi	AC016386 Homo sapi
c 642	33	3.2	175043	2	AC024246	Homo sapi	AC024246 Homo sapi
c 643	33	3.2	175307	2	AC093246	Homo sapi	AC093246 Homo sapi
c 644	33	3.2	176584	2	AC027812	Homo sapi	AC027812 Homo sapi
c 645	33	3.2	177426	2	AC021222	Homo sapi	AC021222 Homo sapi
c 646	33	3.2	178087	9	AC005089	Homo sapi	AC005089 Homo sapi
c 647	33	3.2	178089	2	AC093133	Homo sapi	AC093133 Homo sapi
c 648	33	3.2	178460	2	AL139040	Homo sapi	AL139040 Homo sapi
c 649	33	3.2	179182	2	AC092965	Homo sapi	AC092965 Homo sapi
c 650	33	3.2	179454	2	AC024177	Homo sapi	AC024177 Homo sapi
c 651	33	3.2	180038	9	AC068393	Homo sapi	AC068393 Homo sapi
c 652	33	3.2	180283	9	AF134726	Homo sapi	AF134726 Homo sapi
c 653	33	3.2	180849	2	AC023996	Homo sapi	AC023996 Homo sapi
c 654	33	3.2	180849	2	AC023996	Homo sapi	AC023996 Homo sapi



801	31	3.0	164945	9	AC089987	AC089987 Homo sapi	874	30	2.9	44183	2	AC093229	Homo sapi
c 802	31	3.0	165110	9	AL450311	AL450311 Human DNA	c 875	30	2.9	44183	2	AC093229	Homo sapi
803	31	3.0	166234	9	HS273P12	HS273P12 Human DNA	876	30	2.9	47812	9	AC005351	Homo sapi
804	31	3.0	167025	2	AC079866	AC079866 Homo sapi	c 877	30	2.9	59847	9	AL136112	Human DNA
c 805	31	3.0	167880	9	HU091323	HU091323 Human Chrom	c 878	30	2.9	63609	9	AL133343	Human DNA
c 806	31	3.0	169032	2	AC061965	AC061965 Homo sapi	c 879	30	2.9	66741	9	HS460J8	Human DNA
c 807	31	3.0	169181	2	AC023303	AC023303 Homo sapi	880	30	2.9	80600	9	AL121903	Human DNA
c 808	31	3.0	170896	2	AC004524	AC004524 Homo sapi	881	30	2.9	92460	9	AL136228	Human DNA
c 809	31	3.0	170000	2	AC011010	AC011010 Homo sapi	882	30	2.9	93165	9	AL135794	Human DNA
c 810	31	3.0	170896	2	CNS01D50	AL121857 Human chr	883	30	2.9	94673	2	AC009018	Homo sapi
c 811	31	3.0	171757	9	AL356095	AL356095 Human DNA	c 884	30	2.9	98180	2	AC091867	Homo sapi
c 812	31	3.0	172079	9	AL357449	AL357449 Homo sapi	c 885	30	2.9	98595	9	AP001329	Homo sapi
c 813	31	3.0	172376	2	AL353749	AL353749 Pan trogl	c 886	30	2.9	101389	2	AL131365	Homo sapi
c 814	31	3.0	173109	2	AC091400	AC091400 Pan trogl	c 887	30	2.9	101719	2	AL133365	Homo sapi
c 815	31	3.0	173229	2	AC091504	AC091504 Homo sapi	c 888	30	2.9	102119	9	AL139005	Human DNA
c 816	31	3.0	173341	2	AC021954	AC021954 Homo sapi	c 889	30	2.9	108554	9	AC005226	Homo sapi
c 817	31	3.0	173463	9	AL359092	AL359092 Human DNA	890	30	2.9	114983	2	AC093209	Homo sapi
c 818	31	3.0	173957	2	AC091671	AC091671 Paplo cyn	891	30	2.9	115932	9	HS1107	Human DNA
c 819	31	3.0	175835	2	AC024074	AC024074 Homo sapi	892	30	2.9	117031	9	AC008973	Homo sapi
c 820	31	3.0	176388	2	AC025273	AC025273 Homo sapi	893	30	2.9	119942	2	AC024582	Homo sapi
c 821	31	3.0	176407	2	AP003040	AP003040 Homo sapi	c 894	30	2.9	119942	2	AC024582	Homo sapi
c 822	31	3.0	178045	2	AC052719	AC052719 Homo sapi	c 895	30	2.9	120997	9	AL133005	Human DNA
c 823	31	3.0	178325	2	AC073046	AC073046 Homo sapi	896	30	2.9	121208	9	HS291J10	Human DNA
c 824	31	3.0	178434	2	AL354989	AL354989 Homo sapi	897	30	2.9	122103	9	AL136101	Human DNA
c 825	31	3.0	178904	2	AC021704	AC021704 Homo sapi	898	30	2.9	124281	9	AL139125	Homo sapi
c 826	31	3.0	179974	2	AC026942	AC026942 Homo sapi	899	30	2.9	124437	9	AL139125	Homo sapi
c 827	31	3.0	180044	2	AC026942	AC026942 Homo sapi	900	30	2.9	131692	9	HSJ881L22	Homo sapi
c 828	31	3.0	180942	33	AC019194	AC019194 Homo sapi	901	30	2.9	131841	9	HSJ881L22	Homo sapi
c 829	31	3.0	181829	9	AC023425	AC023425 Homo sapi	c 902	30	2.9	134594	2	AC083822	Homo sapi
c 830	31	3.0	183451	2	AC024348	AC024348 Homo sapi	c 903	30	2.9	135305	2	AC090427	Homo sapi
c 831	31	3.0	183813	2	AC012204	AC012204 Homo sapi	c 904	30	2.9	136129	2	AC092483	Homo sapi
c 832	31	3.0	184637	2	AC012043	AC012043 Homo sapi	c 905	30	2.9	136155	2	AC093346	Homo sapi
c 833	31	3.0	185229	9	AL157762	AL157762 Human DNA	906	30	2.9	139488	9	AC003043	Homo sapi
c 834	31	3.0	185401	2	AC069068	AC069068 Homo sapi	907	30	2.9	139836	2	AC090052	Homo sapi
c 835	31	3.0	185804	2	AC017095	AC017095 Homo sapi	908	30	2.9	140026	9	AC005840	Homo sapi
c 836	31	3.0	185827	9	AC084357	AC084357 Homo sapi	c 909	30	2.9	141049	2	AL354662	Homo sapi
c 837	31	3.0	186527	9	AC018921	AC018921 Homo sapi	c 910	30	2.9	141597	2	AC093223	Homo sapi
c 838	31	3.0	190579	2	AC018736	AC018736 Homo sapi	c 911	30	2.9	142344	9	AL136525	Human DNA
c 839	31	3.0	191905	2	AL512770	AL512770 Homo sapi	c 912	30	2.9	145340	9	AC026539	Homo sapi
c 840	31	3.0	191942	9	AC079951	AC079951 Homo sapi	c 913	30	2.9	145833	2	AC010803	Homo sapi
c 841	31	3.0	192768	9	AC016831	AC016831 Homo sapi	c 914	30	2.9	146487	2	AC022243	Homo sapi
c 842	31	3.0	196000	9	AL354749	AL354749 Human DNA	915	30	2.9	149901	2	AC010330	Homo sapi
c 843	31	3.0	196327	2	AC092429	AC092429 Homo sapi	c 916	30	2.9	149913	2	AC010330	Homo sapi
c 844	31	3.0	197791	2	AC016153	AC016153 Homo sapi	c 917	30	2.9	151845	2	AC008128	Homo sapi
c 845	31	3.0	199380	2	AC0048379	AC0048379 Homo sapi	c 918	30	2.9	151845	2	AC008128	Homo sapi
c 846	31	3.0	200000	2	AC004618	AC004618 Homo sapi	c 919	30	2.9	153087	2	AC017072	Homo sapi
c 847	31	3.0	200536	9	AL356093	AL356093 Human DNA	c 920	30	2.9	153716	2	AC069187	Homo sapi
c 848	31	3.0	200594	9	AC005632	AC005632 Homo sapi	921	30	2.9	154270	2	AC023391	Homo sapi
c 849	31	3.0	201794	9	CNS0180W	AL109759 Human chr	922	30	2.9	154745	2	AC010860	Homo sapi
c 850	31	3.0	204272	2	AC007215	AC007215 Homo sapi	923	30	2.9	154960	2	AC011948	Homo sapi
c 851	31	3.0	210000	2	AC004555	AC004555 Homo sapi	c 924	30	2.9	155459	2	AC016619	Homo sapi
c 852	31	3.0	215727	2	AL139407	AL139407 Homo sapi	c 925	30	2.9	155769	2	AC073250	Homo sapi
c 853	31	3.0	220128	2	AL139407	AL139407 Homo sapi	c 926	30	2.9	156390	2	AL1456	Homo sapi
c 854	31	3.0	236680	2	AC024120	AC024120 Homo sapi	c 927	30	2.9	157058	2	CNS05TBY	Homo sapi
c 855	31	3.0	237044	2	AP000926	AP000926 Homo sapi	c 928	30	2.9	157186	2	AL592076	Homo sapi
c 856	31	3.0	247473	2	AC011137	AC011137 Homo sapi	c 929	30	2.9	159057	2	AC024705	Homo sapi
c 857	31	3.0	340000	9	AP001725	AP001725 Homo sapi	c 930	30	2.9	160209	9	AC012324	Homo sapi
c 858	30	2.9	1637	9	AF349314	AF349314 Sequence	c 931	30	2.9	161198	2	AC015867	Homo sapi
c 859	30	2.9	1762	6	AK147248	AK147248 Sequence	c 932	30	2.9	162466	2	AC007011	Homo sapi
c 860	30	2.9	1804	9	AK022325	AK022325 Homo sapi	c 933	30	2.9	162575	9	AL157935	Homo sapi
c 861	30	2.9	1916	9	BC008304	BC008304 Homo sapi	c 934	30	2.9	162575	9	AL157935	Homo sapi
c 862	30	2.9	1994	9	AC092308	AC092308 Homo sapi	c 935	30	2.9	162575	9	AL157935	Homo sapi
c 863	30	2.9	12845	2	AC092308	AC092308 Homo sapi	c 936	30	2.9	162575	9	AL157935	Homo sapi
c 864	30	2.9	14402	2	AC091668	AC091668 Homo sapi	c 937	30	2.9	164241	2	AL157935	Homo sapi
c 865	30	2.9	17982	2	AC093239	AC093239 Homo sapi	c 938	30	2.9	164628	2	AC073646	Homo sapi
c 866	30	2.9	19579	9	AC007487	AC007487 Homo sapi	c 939	30	2.9	164628	2	AC073646	Homo sapi
c 867	30	2.9	29579	9	AC000122	AC000122 Homo sapi	c 940	30	2.9	164628	2	AC073646	Homo sapi
c 868	30	2.9	32260	2	AC093237	AC093237 Homo sapi	941	30	2.9	164884	2	AC020909	Homo sapi
c 869	30	2.9	32260	2	AC093237	AC093237 Homo sapi	942	30	2.9	164884	2	AC020909	Homo sapi
c 870	30	2.9	32260	2	AL589656	AL589656 Human DNA	943	30	2.9	166264	2	AL330864	Homo sapi
c 871	30	2.9	39383	9	AC022149	AC022149 Homo sapi	944	30	2.9	166264	2	AL330864	Homo sapi
c 872	30	2.9	43390	9	AC007191	AC007191 Homo sapi	c 945	30	2.9	166915	2	AL330864	Homo sapi
c 873	30	2.9	43390	9	AC007191	AC007191 Homo sapi	c 946	30	2.9	166915	2	AL330864	Homo sapi

947	30	2.9 167074	2	AC026643	
948	30	2.9 167357	9	AL138976	Human chr
949	30	2.9 167408	2	AC023020	
c 950	30	2.9 168149	2	AL158843	
951	30	2.9 168230	2	AC010624	
c 952	30	2.9 168230	2	AC010624	
953	30	2.9 169323	2	AL161906	
954	30	2.9 169343	9	AC016955	
955	30	2.9 169623	2	AC055764	
956	30	2.9 169845	2	AC019054	
957	30	2.9 170114	2	AC016143	
958	30	2.9 170470	2	AC022696	
959	30	2.9 170655	2	AP001012	
c 960	30	2.9 170655	2	AP001012	
961	30	2.9 170954	2	AL356740	
c 962	30	2.9 171189	2	AC008966	
963	30	2.9 171788	9	AC073089	
964	30	2.9 171922	2	AC091045	
c 965	30	2.9 172601	2	AC021455	
c 966	30	2.9 173149	2	AL391218	
c 967	30	2.9 173213	2	AC066692	
968	30	2.9 173452	9	AC003013	
969	30	2.9 173575	2	AC016932	
c 970	30	2.9 173646	2	AC016542	
c 971	30	2.9 174466	2	AC068768	
c 972	30	2.9 174476	2	AC037687	
c 973	30	2.9 174480	2	AC026069	
974	30	2.9 174579	2	AC016666	
c 975	30	2.9 175271	2	AC090698	
c 976	30	2.9 175670	2	AL358353	
c 977	30	2.9 175789	2	AP000425	
c 978	30	2.9 175999	2	AC021154	
979	30	2.9 176391	9	AL365400	
980	30	2.9 176571	2	AC027111	
981	30	2.9 177066	2	AC069257	
c 982	30	2.9 177289	2	AC090176	
c 983	30	2.9 178246	2	AL591506	
984	30	2.9 179134	2	AC022817	
985	30	2.9 180486	2	AC007621	
986	30	2.9 180520	2	AC019152	
c 987	30	2.9 180522	2	CNS01DSH	
988	30	2.9 182018	2	AC015464	
c 989	30	2.9 182047	2	AC013462	
c 990	30	2.9 182214	2	AC067869	
991	30	2.9 182414	2	AC055877	
992	30	2.9 182747	2	AL512330	
993	30	2.9 183842	2	AC080084	
994	30	2.9 184476	2	AL138750	
995	30	2.9 184853	2	AC020723	
996	30	2.9 186607	2	AC092981	
997	30	2.9 187760	2	AC092755	
c 998	30	2.9 188076	2	AC023629	
c 999	30	2.9 188249	2	AL450023	
1000	30	2.9 188346	2	AC010189	

ALIGNMENTS

RESULT	1
AX107720	
LOCUS	AX107720 1041 bp
DEFINITION	Sequence 9 from Patent WO0123579.
ACCESSION	AX107720
VERSION	AX107720.1
KEYWORDS	GR:13923201
SOURCE	human.
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 1041).  
Donoho G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.

TITLE	Human kinase proteins and polynucleotides encoding the same
JOURNAL	Patent: WO 0123579-A 9 05-APR-2001;
FEATURES	Lexicon Genetics Incorporated (US)
source	Location/Qualifiers
	1..1041
BASE COUNT	/organism="Homo sapiens"
ORIGIN	/db_xref="taxon:9606"
	344 a 190 c 228 g 279 t
Query Match	100.0%; Score 1041; DB 6; Length 1041;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1041:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 atgaaaagatataaaaaaataagtaagactggaagggctcttatgggtgtgattcaaa 60
Db	1 ATGGAAAGATATGAAAATTAAGCTAAGACTGGAGAGGCTCTATGGGTGTGATTCMAA 60
QY	61 tgcagaacaaaacctctggacaagttagctgttaaaaaatttggaattcgaatcgaagat 120
Db	61 TGCAGAAACAAAACCTCTGGCAAGTAGTAGCTGTAAAAAATTTGTGGAATCTGAAGAT 120
QY	121 gatcctgtttaaagaaataagcactaagagaaatacgtatgttgaagcaattaaacat 180
Db	121 GATCCTGTCTTAAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180
QY	181 ccaaatcttgtaacctcatcgagggtgttcaggagaaaaagaaaatgcattagtcttt 240
Db	181 CCAAACTTGTGAACCTCATCGAGGTGTTCAAGAGAAAAGAAAATGCATTAGTCTTTT 240
QY	241 gaactactgtatcatacaacttttaaatgagctggaagaaacccaatggaattgctgat 300
Db	241 GAATACTGTGATCATACACTTTTAATGAGCTGGAGAAACCCAAATGAGTTCGTGAT 300
QY	301 ggaagtatacaaaagcgtattatggcaaaccttcaagctcttaattctgcatatcat 360
Db	301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
QY	361 aactgtattcacagagatataaaacctgaaatatttcaataactaaagaaagaaatac 420
Db	361 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATACTAACAAGGAATATTC 420
QY	421 aagatttgcacttcgggtttgcacaaattctgattccaggagatgcctacaccgattat 480
Db	421 AAGATTGTGACTTCGGGTTTGCACAAATCTGTATCCAGGAGATGCCATACCCGATTAT 480
QY	481 gtagctacagatggtaaccgagctcctgaaactcttctgtgggagatactcagatggttct 540
Db	481 GTAGCTACGATGGTATACCGAGCTCCTGAACCTCTTGTGGGAGATACCTCAGTATGTTCT 540
QY	541 tcagtcgatatatgggtattggttgggttttttcgagagctcctgacaggccagccatg 600
Db	541 TCAGTCGATATATGGGCTATTGGTTGTTGTTTTGCGAGAGCTCTGACAGCCGACCTG 600
QY	601 tggcctgaaatacagatggaccactttatctgataatacagaacactagtagagacg 660
Db	601 TGGCCTGGAAATAACAGATGTGACCAACTTTATCTGATAATACAGAACACTAGTAGAGCG 660
QY	661 ggggttcgcatgttgacaggctgggtctcgaaactcttgacgtcgaagtgatccactgcc 720
Db	661 GGGTTTCGCATGTTGACAGGCTGGTCTCGAACTCTTGAACCTTTCAGTGTATCCACTGCC 720
QY	721 gtacacctcaagtgctggaattacaggaataattatcccaagacatcaatcatcttt 780
Db	721 GTAGCCCTCAAGTGTCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780
QY	781 aagaagtaacgggtttttccatggtcagtcagtatcacctgagccagaagacatgaaactctt 840
Db	781 AAAAGTAACGGGTTTTTCCATGTCATCAGTATACCTGACGACAGACATGGAACACTTT 840
QY	841 gaggaagaaagttctcagatgttctcctctggtcctggaactctgaactcatgaaggggtgctgaag 900

Db	841	GAGGAAAAAGTTCTCAGATGTTTCATCTCTGGCTCTGCACTTCATGCAAGGGGTCTCTGAAG	900
Qy	901	atgaatcaagatagatgaatcaactgttcccaactccttgagagctcctacttattgattct	960
Db	901	ATGAATCCAGATGACAGATTACCTGTTCCTCCAACTCCTGGAGAGCTCCTATTGATTCT	960
Qy	961	tttcaagaggcccaaatataaagaagaacgcgttaataaggaagaagaacagaaagaccccaa	1020
Db	961	TTTCAAGAGGCCCAAAATTAAGAAGAAAGCAGCTAATGAAGGAGAAACAGAACGCCAA	1020
Qy	1021	caagtacttccgtccaaagt	1041
Db	1021	CAGGTACTTCCGCTCAAAAGT	1041
RESULT 2			
LOCUS	AX107714	1068 bp	DNA
DEFINITION	Sequence 3 from Patent WO0123579.		PAT
ACCESSION	AX107714		
VERSION	AX107714.1	GI:13923198	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 1068)		
TITLE	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.		
JOURNAL	Human Kinase proteins and polynucleotides encoding the same		
FEATURES	Patent: WO 0123579-A 3 05-APR-2001; Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers		
	1..1068		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	344 a 197 c 240 g 287 t		
ORIGIN			
Query Match 98.5%; Score 1025; DB 6; Length 1068;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	atggaataagatgaaaaattagctaaagactggaagaagggtcttatgggttgattcaaa	60
Db	1	ATGGAATAAGTATGAAAAATTAGCTAAGACTTGAAGAGGTCTTATGGGTGTGATTCAAA	60
Qy	61	tcgagaacaaacctctcgacaagtagtagctgttaaaaaatttggaaatctgaagat	120
Db	61	TGCAGAAACAAAGCTCTGGACAAGTAGTCTGTAAAAAATTGTGGAAATCTGAAGAT	120
Qy	121	gatcctgttgaagaataagcactaagaagaataacgtatgtgaagcaatttaaacat	180
Db	121	GATCCTGTTGTTAGAAATAGCAGCTAAGAGAAATACGTATGTGAAGCAATTAACAT	180
Qy	181	ccaaattctgtgaacctcagcaggtgttcagagaagaagaagaagcatttagtttt	240
Db	181	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGAGAGAAAGGAAATGCAATTGATTTT	240
Qy	241	gaatcgtgtatcacacttttaaatgagctggaagaacccaatggagttgctgat	300
Db	241	GAATCTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCAATGGAGTTGCTGAT	300
Qy	301	ggagatcaaaagcgtattatgcgaacacttcaagctcttaatttctgtcatatacat	360
Db	301	GGAGTATCAAAAGCGTATTATGGCAACACTTCAAGCTTTAATTCTGTCATATACAT	360
Qy	361	aactgtattcacagagatatataaacctgaaaaattcttaataacaaagcaagaataac	420
Db	361	AACGTATTACAGAGATATAAAACCTGAAAAATTATTCATAACTAGCAGGAATATC	420
Qy	421	aagattgtgacttcgggtttgcaacaaattctgtattccaggagatgctacacccattat	480

Db	421	AAGATTTGTGACATCGGGTTTGACAAAAATTCGATTTCAGGAGATGCTACACCGATTAT	480
Qy	481	gtagctacagataggtaccagagctcctgaactcttggtgggagatactcaatgattct	540
Db	481	GTAGCTACGAGATGGTACCAGAGCTCCTGAACCTTCTTGTGGGAGATACTAGTATGGTTCT	540
Qy	541	tcagtcatatatgggctattggttggttttgcagagctcctgcagagccagccactg	600
Db	541	TCAGTCGATATATGGGCTATTGTTGTTGTTTTCAGAGCTCCTGACAGCCAGCCACTG	600
Qy	601	tggcctggaaaaatcagatgtggacaaactttatctaatcagaacactagtagagacg	660
Db	601	TGGCCTGGAAAAATCAGATGTGGACCAACTTTTATCTGATTAATCAGAACACTAGTAGAGCG	660
Qy	661	gggtttccgcatgttgaccaggtggtctcgaactcttgacgtcaagtgaatccacctgc	720
Db	661	GGGTTTCGCCATGTTGACCAGGCTGCTCGAACTCTTGACGTCAAGTGATCCACCTGCC	720
Qy	721	gtagccttcaagtgcgtgaattacagaaaaataatcccaagacatacaatcaatttt	780
Db	721	GTAGCCTCTCAAGTGTCTGGAATTACAGGAAATTAATCCCAAGACATCAATCAATCTTT	780
Qy	781	aaaagtaacgggttttccatggcatcagtataccctgagccagacagacatgaaactctt	840
Db	781	AAAAGTAACGGGTTTTTCCATGGCATCAGTATACCTGAGCCAGACATGGAACACTTT	840
Qy	841	gagaaaaattctcagatgttcctctggtcgtgaaactctcatgaaggggtgtctgaag	900
Db	841	GAGAAAAATTCTCAGATGTTTCCTCTGTGGCTCTGAACCTTCATGAAGGGGTGCTGAAG	900
Qy	901	atgaatccagatgacagattaaacctgttcccaactctggagagctcctactttgattct	960
Db	901	ATGAATCCAGATGACAGATTAACTGTTCCTCAACTCTCTGGAGAGCTCTACTTTGATTCT	960
Qy	961	tttcaagaggcccaaatataaagaaaagcagcgttaataaggaagaagaagaagacccaa	1020
Db	961	TTTCAAGAGGCCCAAAATTAAGAAGAAAGCAGCTAATGAAGGAGAAACAGAGACGCCAA	1020
Qy	1021	caaggt	1025
Db	1021	CAGGT	1025
RESULT 3			
LOCUS	AX107722	945 bp	DNA
DEFINITION	Sequence 11 from Patent WO0123579.		PAT
ACCESSION	AX107722		
VERSION	AX107722.1	GI:13923202	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.		
TITLE	Human Kinase proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0123579-A 11 05-APR-2001; Lexicon Genetics Incorporated (US)		
FEATURES	Location/Qualifiers		
	1..945		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	323 a 165 c 202 g 255 t		
ORIGIN			
Query Match 62.68%; Score 652; DB 6; Length 945;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

Qy 1 atggaagaatgataaaattagctaaagactggaaggggtcttatggggtgtattcaaa 60  
 Db 1 ATGGAAAGATATGAAAAATTAGCTAAGACTGGAGAAGGCTCTTATGGGGTTGTATTCAA 60  
 Qy 61 tgcagaacaaacctctggacaagtagctgttataaaatttgggaattctgaagat 120  
 Db 61 TGCAGAACAAACCTCTGGACAAGTAGTACTGTTAAAAAATTTGTGGAATCTGAAGAT 120  
 Qy 121 gatcctgttgaagaataagcactaagagaatacgtatgttgaagaataaacaat 180  
 Db 121 GATCCTGTTTGAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAAT 180  
 Qy 181 ccaaatcttgaaccctcatcagaggttcaagagaaagaaagaaatgcatttaatttt 240  
 Db 181 CCAAACTTGTGACCTCATCGAGGTGTTGAGGAGAAAGAGAAATGCAATTAGTTTTT 240  
 Qy 241 gaatactgtatcatatacattttaaatgagctggagaagaaacccaatggagttcgtgat 300  
 Db 241 GAATACTGTATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300  
 Qy 301 ggaagtcatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360  
 Db 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCTATATACAT 360  
 Qy 361 aactgtattcacagagataaaacctgaaatatttctaataactaagcaaggaataatc 420  
 Db 361 AACTGTATTACAGAGATATAAACCTGAAATATTCTTAATAAGCAAGGAATAATC 420  
 Qy 421 aagattgtacttcgggtttgcacaaattctgtaccagagagatgctacaccgattat 480  
 Db 421 AAGATTGTACTTCGGGTTGGACAAATCTGATTCAGAGAGATGCTTACACCGATTAT 480  
 Qy 481 gtaactacagagatggtaaccagactcctgaactctctgtggagatactcaatagttct 540  
 Db 481 GTACTACAGAGATGACCGAGCTCTGAACCTCTTGTGGAGATCTACGATATGGTTCT 540  
 Qy 541 tcaatgatataatgggtattggtgttttttgcagagctcctgcagagccagccactg 600  
 Db 541 TCAGTCGATATATGGCTATTGTTGTTGTTTTTGCAGAGCTCCTGCAGAGCCAGCCACTG 600  
 Qy 601 tggcctggaataatcagatgtggaccaactttatctatgataatcagaacactag 652  
 Db 601 TGGCTGGAAATCAGATGTGGACCAACTTTATCTGATTAATCAGAACACTAG 652  
 RESULT 4  
 AX107716 972 bp DNA PAT 30-APR-2001  
 LOCUS AX107716 Sequence 5 from Patent WO0123579.  
 DEFINITION AX107716.1 GI:13923199  
 ACCESSION AX107716  
 VERSION AX107716.1  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 972)  
 AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.  
 TITLE Human kinase proteins and polynucleotides encoding the same  
 JOURNAL Patent: WO 0123579-A 5 05-APR-2001;  
 Lexicon Genetics Incorporated (US)  
 FEATURES Location/Qualifiers  
 source 1..972  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 323 a 172 c 214 g 263 t  
 ORIGIN

Query Match 62.6%; Score 652; DB 6; Length 972;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggaagaatgataaaattagctaaagactggaaggggtcttatggggtgtattcaaa 60  
 Db 1 ATGGAAAGATATGAAAAATTAGCTAAGACTGGAGAAGGCTCTTATGGGGTTGTATTCAA 60  
 Qy 61 tgcagaacaaacctctggacaagtagctgttataaaatttgggaattctgaagat 120  
 Db 61 TGCAGAACAAACCTCTGGACAAGTAGTACTGTTAAAAAATTTGTGGAATCTGAAGAT 120  
 Qy 121 gatcctgttgaagaataagcactaagagaatacgtatgttgaagaataaacaat 180  
 Db 121 GATCCTGTTTGAAGAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAAT 180  
 Qy 181 ccaaatcttgaaccctcatcagaggttcaagagaaagaaagaaatgcatttaatttt 240  
 Db 181 CCAAACTTGTGACCTCATCGAGGTGTTGAGGAGAAAGAGAAATGCAATTAGTTTTT 240  
 Qy 241 gaatactgtatcatatacattttaaatgagctggagaagaaacccaatggagttcgtgat 300  
 Db 241 GAATACTGTATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300  
 Qy 301 ggaagtcatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360  
 Db 301 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCTATATACAT 360  
 Qy 361 aactgtattcacagagataaaacctgaaatatttctaataactaagcaaggaataatc 420  
 Db 361 AACTGTATTACAGAGATATAAACCTGAAATATTCTTAATAAGCAAGGAATAATC 420  
 Qy 421 aagattgtacttcgggtttgcacaaattctgtaccagagagatgctacaccgattat 480  
 Db 421 AAGATTGTACTTCGGGTTGGACAAATCTGATTCAGAGAGATGCTTACACCGATTAT 480  
 Qy 481 gtaactacagagatggtaaccagactcctgaactctctgtggagatactcaatagttct 540  
 Db 481 GTACTACAGAGATGACCGAGCTCTGAACCTCTTGTGGAGATCTACGATATGGTTCT 540  
 Qy 541 tcaatgatataatgggtattggtgttttttgcagagctcctgcagagccagccactg 600  
 Db 541 TCAGTCGATATATGGCTATTGTTGTTGTTTTTGCAGAGCTCCTGCAGAGCCAGCCACTG 600  
 Qy 601 tggcctggaataatcagatgtggaccaactttatctatgataatcagaacactag 652  
 Db 601 TGGCTGGAAATCAGATGTGGACCAACTTTATCTGATTAATCAGAACACTAG 652  
 RESULT 5  
 AX107724 1819 bp DNA PAT 30-APR-2001  
 LOCUS AX107724 Sequence 13 from Patent WO0123579.  
 DEFINITION AX107724  
 ACCESSION AX107724  
 VERSION AX107724.1 GI:13923203  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1819)  
 AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.  
 TITLE Human kinase proteins and polynucleotides encoding the same  
 JOURNAL Patent: WO 0123579-A 13 05-APR-2001;  
 Lexicon Genetics Incorporated (US)  
 FEATURES Location/Qualifiers  
 source 1..1819  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 539 a 368 c 406 g 505 t  
 ORIGIN

Query Match 54.9%; Score 572; DB 6; Length 1819;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-306;

Matches		572;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
y	454	attcagagagatgctacacgattatgtagtacgagatggtaccgagctcctgaactt	513							
b	897	ATTCCAGGAGATGCTACACCGATTATGTAGCTACGAGATGGTACCGAGCTCTGAAC	956							
y	514	cttgtgggagatactcagatggtttcttcagtcagtcgatatatggctattggtgtgtttt	573							
b	957	CTTGTGGGAGATCTCAGTATGGTTCCTTCAGTCGATATATGGCTATTGGTGTGTTTT	1016							
y	574	gcaagctcctgacagccagccactgtggcctggaaatacagatgagaccactttat	633							
b	1017	GCAGAGCTCTGACAGGCCAGCCACTGTGGCTGGAAATCAGATGTGACCAACTTAT	1076							
y	634	ctgataatcagaacactagtagacggggttttcgcatgttgaccagctggtctcgaa	693							
b	1077	CTGATANTCAGAACACTAGTAGAGCGGGGTTTCGCATGTTGACCCAGGCTGGTCGAA	1136							
y	694	ctcttgacatgaagtatccacctgcccgtagcctctcaagtgctggaattacagaaaa	753							
b	1137	CTCTTGACGTCAAGTGATCCACCTGCCGTAGCCTCTCAAGTCTGGAATTCAGAGAAA	1196							
y	754	ttaatcccaagacatcaatcaatctttaaaagtaacgggtttttccatggcatcagata	813							
b	1197	TTAATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTCATGGCATCAATA	1256							
y	814	cctgagccagagacatggaaacctcttgagaaagttctcagatgttccatcctgtggt	873							
b	1257	CCTGAGCCAGAGACATGGAACCTCTTGAGGAAAGTCTCAGATGTTTCATCTGTGCT	1316							
y	874	ctgaactctgaagggtgtctgaagatgaatccagatgacagattaacctgttcccaa	933							
b	1317	CTGAACCTTCATGAAGGGTGTCTGAAGATGAATCCAGATGACAGATTAACTGTTC	1376							
y	934	ctcctgagagctcctacttgattcttccaagagggcccaataaagaaagacagt	993							
b	1377	CTCCTGGAGAGCTCTCTACTTTCATCTTTTCAAGAGGCCCAATTTAAAGAAAGC	1436							
y	994	aatgaaggaagaacagaagcgcacaggt	1025							
b	1437	AATGAAGGAAGAACAAGAGCGCCACAGGT	1468							
RESULT 6										
LOCUS	AX107712	561 bp	DNA	PAT	30-APR-2001					
DEFINITION	Sequence 1 from Patent W00123579.									
ACCESSION	AX107712									
VERSION	AX107712.1 GI:13923197									
KEYWORDS	human.									
SOURCE	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.									
TITLE	Human kinase proteins and polynucleotides encoding the same									
JOURNAL	Patent: WO 0123579-A 1 05-APR-2001;									
FEATURES	Lexicon Genetics Incorporated (US)									
source	Location/Qualifiers									
	1. .561									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
BASE COUNT	189 a 87 c 112 g 173 t									
ORIGIN										
Query Match 43.6%; Score 454; DB 6; Length 561;										
Best Local Similarity 100.0%; Pred. No. 8.2e-241;										
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
Qy	1	atggaagatgaaaaattagctaagactggagaagggtcttattggtgttatttcaaa	60							

Db	1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
----	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

QY 181 ccaaatctgtgaacctcatcgagggtgttcaggagaaagaaagcaatgcatgttttt 240  
Db 181 CCAAACTTGTGACCTCATCGAGGTGTTACAGAGAAAGAAATGCAUATTAGTTTTT 240  
QY 241 gaatactgtatcatcacattttaaatgagctggaagaacacccaaatggagttgctgat 300  
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAACCCAAATGGAGTTGCTGAT 300  
QY 301 ggaagtcatcaaaagcgtattatgcaaacacttcaagcttcttaattctgtcatatacat 360  
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAAATTTCTGTCATATACAT 360  
QY 361 aactgtattcacagagataaaacctgaaataattcttaactaaagcaagaataatc 420  
Db 361 AACTGTATTTCACAGAGATATAAACTGAAATATTCTTAATACTAAGCAAGCAATATC 420  
QY 421 aagatttgacttcgggtttgcacaaattctga 454  
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATTCGA 454  
RESULT 8  
AX166534  
LOCUS AX166534 1083 bp DNA PAT 22-JUN-2001  
DEFINITION Sequence 25 from Patent WO0138503.  
ACCESSION AX166534  
VERSION AX166534.1 GI:14546879  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1083)  
AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,  
TITLE Novel human protein kinases and protein kinase-like enzymes  
JOURNAL Patent: WO 0138503-A 25 31-MAY-2001,  
Sugen, Inc. (US)  
FEATURES  
Source Location/Qualifiers  
1..1083  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 366 a 199 c 225 g 293 t  
ORIGIN  
Query Match 43.6%; Score 454; DB 6; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 8.3e-241;  
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atggaagagatgaaattagctgaagactggagaagggtcttatgggtgttatcaaa 60  
Db 1 ATGGAAGATGTAAGAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAA 60  
QY 61 tgagaaacaaacctctggacaagttagctgaagactggaagactggaagactggaagact 120  
Db 61 TGAGAAACAAACCTCTGACAACTGACGCTGTTAAAGAAATTTGGGAATCTGAAGAT 120  
QY 121 gatcctgttgaagaaataagcactaagaagaatacgtatgttgaagcaattaaacaat 180  
Db 121 GATCCTGTTGTAGAAATATGACCTAAGAGAAATACGTATGTTGAAGCAATTTAAACAT 180  
QY 181 ccaaatctgtgaacctcatcgagggttcaggagaaagaaagcaattgatttttt 240  
Db 181 CCAAACTCTGTGAACCTCATCGAGGTGTTACAGAGAAAGAAATGCAUATTAGTTTTT 240  
QY 241 gaatactgtatcatcacattttaaatgagctggaagaacacccaaatggagttgctgat 300  
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAACCCAAATGGAGTTGCTGAT 300  
QY 301 ggaagtcatcaaaagcgtattatgcaaacacttcaagcttcttaattctgtcatatacat 360  
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAAATTTCTGTCATATACAT 360

QY 361 aactgtattcacagagataaaacctgaaataattcttaactaaagcaagaataatc 420  
Db 361 AACTGTATTTCACAGAGATATAAACTGAAATATTCTTAATACTAAGCAAGCAATATC 420  
QY 421 aagatttgacttcgggtttgcacaaattctga 454  
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATTCGA 454  
RESULT 9  
AX056404  
LOCUS AX056404 911 bp DNA PAT 13-JAN-2001  
DEFINITION Sequence 48 from Patent WO0073469.  
ACCESSION AX056404  
VERSION AX056404.1 GI:12229111  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 911)  
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudarsanam,S.  
TITLE Protein kinases  
JOURNAL Patent: WO 0073469-A 48 07-DEC-2000;  
Sugen, Inc. (US)  
FEATURES  
Source Location/Qualifiers  
1..911  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 293 a 182 c 178 g 258 t  
ORIGIN  
Query Match 35.4%; Score 368; DB 6; Length 911;  
Best Local Similarity 100.0%; Pred. No. 5.3e-193;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 285 aaatggagttgctgatggatgatcaaaagcgtattatggcaaacacttcaagctcttaa 344  
Db 1 AAATGGAGTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAA 60  
QY 345 ttctgtcatatcacataaactgtattcacagagatataaaacctgaaaaatattctataac 404  
Db 61 TTTCTGTCATATACATAAATCTGATTCACAGAGATATAAACTGAAAAATATTCTAATAAC 120  
QY 405 taagcaagaataacaaagatttgcacttcgggtttgcacaaattctgtatccagaga 464  
Db 121 TAAGCAAGGAATAATCAAGATTGTTGACTTCGGGTTTGCACAAATTCATGATCCAGGAGA 180  
QY 465 tgcctacaccgattatgtagctacagagatggtaccgagctcctgaacttcttgggaga 524  
Db 181 TGCTTACACCGATTATGTTAGCTACAGAGATGGTACCGAGCTCTGAACTTCTTGTGGAGA 240  
QY 525 tactcagttatggttcttcagtcgatatatgggctattggttggtttttttgcagagctct 584  
Db 241 TACTCAGTATGGTCTTCTCAGTCGATATATGGGCTATTGTTGTTGTTTTCAGAGCTCCT 300  
QY 585 gacagggccagccactgtgacctggaaatcagatggaaccaactttatctgataatcag 644  
Db 301 GACAGGCCAGCCACTGTGGCCTGGAAATACAGATGTGGACCAACTTTTCTGTATATCAG 360  
QY 645 aacactag 652  
Db 361 AACACTAG 368  
RESULT 10  
AC079615  
LOCUS AC079615 80362 bp DNA PRI 03-JUL-2001  
DEFINITION Homo sapiens clone Rpl11-462L6, complete sequence.  
ACCESSION AC079615  
VERSION AC079615.6 GI:14589743



KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS Waterston, R.H. f Homo sapiens clone  
TITLE The sequence of Homo sapiens clone  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 80362)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (04-SEP-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
3 (bases 1 to 80362)  
Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (03-JUL-2001) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Jul 3, 2001 this sequence version replaced gi:14018142.

COMMENT  
----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H.NH0462L06  
----- Location/Qualifiers -----  
1. 80362  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-462L6"  
BASE COUNT 22997 a 15643 c 16118 g 25604 t  
ORIGIN

Query Match 19.0%; Score 198; DB 9; Length 80362;  
Best Local Similarity 100.0%; Pred.No.1.8e-98;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 455 ttccaggagatgcctacaccgattatgtagctacgagatggtaccgagctcgaacttc 514  
Db 73475 TTCCAGGAGATGCCTACACCGATTATGTTAGCTACGAGATGGTACCGAGCTCTGAACCTTC 73534  
Qy 515 ttgtggagatactcagatggtgtttcagtcgatatatgggctattgtgtgttttg 574  
Db 73535 TTGTGGAGATACACTAGTATGGTCTTCAGTCGATATATGGGCTATGGTGTGTTTGT 73594  
Qy 575 cadagctcctgacagggccagccactgtggcctggaataatcagatgtggaccaactttatc 634  
Db 73595 CAGAGCTCTGACAGGGCCAGCCACTGTGGCTGGAAAAATCAGATGTGGACCAACTTTATC 73654  
Qy 635 tgataatcagaacactag 652  
Db 73655 TGATAATCAGAACACTAG 73672

RESULT 11  
AC037489  
ID AC037489 standard; DNA; HTG; 155028 BP.  
XX AC037489;  
AC AC037489;  
XX AC037489.2  
SV AC037489.2  
XX  
XX 10-APR-2000 (Rel. 63, Created)  
DT 14-MAY-2000 (Rel. 63, Last updated, Version 2)  
XX  
XX Homo sapiens chromosome 2 clone RP11-603F24 map 2, WORKING DRAFT SEQUENCE,  
DE 26 unordered pieces.  
DE

XX KW HTGS\_DRAFT; HTGS\_PHASE1.  
XX  
XX Homo sapiens (human)  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
XX  
XX [1]  
RP 1-155028  
RP Birren B., Linton L., Nusbaum C., Lander E.;  
RL "Homo sapiens chromosome 2, clone RP11-603F24";  
RT Unpublished.  
XX  
XX [2]  
RP 1-155028  
RP Birren B., Linton L., Nusbaum C., Lander E., Abraham H., Allen N.,  
RA Anderson S., Baldwin J., Barna N., Bastien V., Bedalov F., Boguslavsky L.,  
RA Boukhalil B., Brown A., Burkett G., Campopiano A., Castle A., Choepel Y.,  
RA Colangelo M., Collins S., Collymore A., Cooke P., DeArelano K., Dewar K.,  
RA Diaz J.S., Dodge S., Domino M., Doyle M., Ferreira P., Fitzhugh W.,  
RA Gage D., Galagan J., Gardyna S., Ginde S., Goyette M., Graham L.,  
RA Grand-Pierre N., Grant G., Hagos B., Heaford A., Horton L., Howland J.C.,  
RA Iliev I., Johnson R., Jones C., Kann L., Karatas A., Klein J., LaRoque K.,  
RA Lamazares R., Landers T., Lehoczy J., Levine R., Liu C., Liu G.,  
RA Locke K., MacDonald P., Marquis N., McCarthy M., McEwan P., McGurk A.,  
RA McKernan K., McPheeters R., Meldrum J., Meneus L., Minova T., Miranda C.,  
RA Mlenga V., Morrow J., Murphy T., Naylor J., Norman C.H., O'Connor T.,  
RA O'Donnell P., O'Neill D., Oliver T.M., Oliver J., Peterson K., Pierre N.,  
RA Pisani C., Pollara V., Raymond C., Riley R., Rogov P., Rothman D., Roy A.,  
RA Santos R., Schauer S., Severy P., Spencer B., Stange-Thomann N.,  
RA Stojanovic N., Subramanian A., Talamas J., Tesfaye S., Theodore J.,  
RA Tirrell A., Travers M., Trigglio J., Vassiliev H., Viel R., Vo A.,  
RA Wilson B., Wu X., Wyman D., Ye W.J., Young G., Zainoun J., Zimmer A.,  
RA Zody M.;  
RT Submitted (09-APR-2000) to the EMBL/GenBank/DBJ databases.  
RL Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,  
RL Cambridge, MA 02141, USA  
XX  
XX On May 12, 2000 this sequence version replaced gi:7528398.  
CC All repeats were identified using RepeatMasker:  
CC Smit, A.F.A. & Green, P. (1996-1997)  
CC http://ftp.genome.washington.edu/RM/RepeatMasker.html  
CC  
CC ----- Genome Center -----  
CC Center: Whitehead Institute/ MIT Center for Genome Research  
CC Center code: WIBR  
CC Web site: http://www.seq.wi.mit.edu  
CC Contact: sequence.submissions@genome.wi.mit.edu  
CC ----- Project Information -----  
CC Center project name: L9434  
CC Center clone name: 603\_F\_24  
CC ----- Summary Statistics -----  
CC Sequencing vector: M13; M77815; 100% of reads  
CC Chemistry: Dye-terminator Big Dye; 100% of reads  
CC Assembly program: Phrap; version 0.960731  
CC Consensus quality: 142960 bases at least Q40  
CC Consensus quality: 148315 bases at least Q30  
CC Consensus quality: 150779 bases at least Q20  
CC Insert size: 164000; agarose-fp  
CC Insert size: 152528; sum-of-contigs  
CC Quality coverage: 4.4 in Q20 bases; agarose-fp  
CC Quality coverage: 4.7 in Q20 bases; sum-of-contigs  
CC -----  
CC \* NOTE: This is a 'working draft' sequence. It currently  
CC \* consists of 26 contigs. The true order of the pieces  
CC \* is not known and their order in this sequence record is  
CC \* arbitrary. Gaps between the contigs are represented as  
CC \* runs of N, but the exact sizes of the gaps are unknown.  
CC \* This record will be updated with the finished sequence  
CC \* as soon as it is available and the accession number will  
CC \* be preserved.  
CC 1 1104: contig of 1104 bp in length  
CC 1105 1204: gap of 100 bp  
CC

```

CC * 1205 2330: contig of 1126 bp in length
CC * 2331 2430: gap of 100 bp
CC * 2431 4965: contig of 2535 bp in length
CC * 4966 5065: gap of 100 bp
CC * 5066 7346: contig of 2281 bp in length
CC * 7347 7446: gap of 100 bp
CC * 7447 9679: contig of 2233 bp in length
CC * 9680 9779: gap of 100 bp
CC * 9780 11977: contig of 2198 bp in length
CC * 11978 12077: gap of 100 bp
CC * 12078 14038: contig of 1961 bp in length
CC * 14039 14138: gap of 100 bp
CC * 14139 15917: contig of 1779 bp in length
CC * 15918 16017: gap of 100 bp
CC * 16018 18776: contig of 2759 bp in length
CC * 18777 18876: gap of 100 bp
CC * 18877 21090: contig of 2214 bp in length
CC * 21091 21190: gap of 100 bp
CC * 21191 23923: contig of 2733 bp in length
CC * 23924 24023: gap of 100 bp
CC * 24024 27664: contig of 3641 bp in length
CC * 27665 27764: gap of 100 bp
CC * 27765 31535: contig of 3771 bp in length
CC * 31536 31635: gap of 100 bp
CC * 31636 36572: contig of 4937 bp in length
CC * 36573 36672: gap of 100 bp
CC * 36673 43155: contig of 6483 bp in length
CC * 43156 43255: gap of 100 bp
CC * 43256 48604: contig of 5349 bp in length
CC * 48605 48704: gap of 100 bp
CC * 48705 54578: contig of 5874 bp in length
CC * 54579 54678: gap of 100 bp
CC * 54679 61303: contig of 6625 bp in length
CC * 61304 61403: gap of 100 bp
CC * 61404 67988: contig of 6585 bp in length
CC * 67989 68088: gap of 100 bp
CC * 68089 75061: contig of 6973 bp in length
CC * 75062 75161: gap of 100 bp
CC * 75162 84322: contig of 9161 bp in length
CC * 84323 84422: gap of 100 bp
CC * 84423 93445: contig of 8923 bp in length
CC * 93446 93445: gap of 100 bp
CC * 93446 107415: contig of 13970 bp in length
CC * 107416 107515: gap of 100 bp
CC * 107516 119348: contig of 11833 bp in length
CC * 119349 119448: gap of 100 bp
CC * 119449 134591: contig of 15143 bp in length
CC * 134592 134691: gap of 100 bp
CC * 134692 155028: contig of 20337 bp in length.
XX
FH key Location/Qualifiers
FH source
FH 1. .155028
FH /chromosome="2"
FH /db_xref="taxon:9606"
FH /organism="Homo sapiens"
FH /map="2"
FH /clone="RP11-603P24"
FH /clone_lib="RPCI-11 Human Male BAC"
FH 1. .1104
FH misc_feature
FH /note="assembly_fragment"
FH misc_feature
FH 1205. .2330
FH /note="assembly_fragment"
FH misc_feature
FH 2431. .4965
FH /note="assembly_fragment"
FH misc_feature
FH 5066. .7346
FH /note="assembly_fragment"
FH misc_feature
FH 7447. .9679
FH /note="assembly_fragment"
FH misc_feature
FH 9780. .11977
FH /note="assembly_fragment"
FH misc_feature
FH 12078. .14038
FH /note="assembly_fragment"

```

```

FT misc_feature
FT 14139. .15917
FT /note="assembly_fragment"
FT misc_feature
FT 16018. .18776
FT /note="assembly_fragment"
FT misc_feature
FT 18877. .21090
FT /note="assembly_fragment"
FT misc_feature
FT 21191. .23923
FT /note="assembly_fragment"
FT misc_feature
FT 24024. .27664
FT /note="assembly_fragment"
FT misc_feature
FT 27765. .31535
FT /note="assembly_fragment"
FT misc_feature
FT 31636. .36572
FT /note="assembly_fragment"
FT misc_feature
FT 36673. .43155
FT /note="assembly_fragment"
FT misc_feature
FT 43256. .48604
FT /note="assembly_fragment"
FT misc_feature
FT 48705. .54578
FT /note="assembly_fragment"
FT misc_feature
FT 54679. .61303
FT /note="assembly_fragment"
FT misc_feature
FT 61404. .67988
FT /note="assembly_fragment"
FT misc_feature
FT 68089. .75061
FT /note="assembly_fragment"
FT misc_feature
FT 75162. .84322
FT /note="assembly_fragment"
FT misc_feature
FT 84423. .93345
FT /note="assembly_fragment"
FT misc_feature
FT 93446. .107415
FT /note="assembly_fragment clone_end:T7 vector_side:right"
FT misc_feature
FT 107516. .119348
FT /note="assembly_fragment"
FT misc_feature
FT 119449. .134591
FT /note="assembly_fragment clone_end:SP6 vector_side:left"
FT misc_feature
FT 134692. .155028
FT /note="assembly_fragment"
XX
SQ Sequence 155028 BP; 43370 A; 31237 C; 31433 G; 46482 T; 2506 other;

Query Match
Best Local Similarity 19.0%; Score 198; DB 33; Length 155028;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 455 ttccaggagatgcctacacgattatgtactagctagctgagctggtaccgagctcctgaacttc 514
Db 45126 TTCCAGGAGATGCTACCGGATTATGTAGTACGAGATGTTACGAGCTCTGAACATTC 45185
QY 515 ttgtgggagatactcagtagtggcttcctcagtcgatatatgggctatggtgtgttttg 574
Db 45186 TTGTGGGAGATCTCAGTATGTTCTTCAGTCCATATATGGCTATATGGCTATGTTGTTTGG 45245
QY 575 cagaactcctgacagccagccactgtgctctgctggaataatcagatgtggaccactttatc 634
Db 45246 CAGAGCTCTTCAGACAGCCAGCCACTGGCCCTGGAAAAATCAGATGTGGACCACTTTATC 45305
QY 635 tgataatcagaacactag 652
Db 45306 TCATAATCAGAACACTAG 45323

```

```

RESULT 12
AC092672
LOCUS
DEFINITION Homo sapiens chromosome 2 clone RP11-603P24, WORKING DRAFT
ACCESSION AC092672
VERSION AC092672.1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT;
SOURCE human.
ORGANISM Homo sapiens

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 164281)  
Waterston,R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 164281)  
Waterston,R.H.  
Direct Submission  
Submitted (19-JUL-2001) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Jul 19, 2001 this sequence version replaced gi:7770604.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H\_NH0603F24  
Drafting center: WIBR  
----- Summary Statistics -----  
Sequencing vector: M13; 47%  
Sequencing vector: plasmid; 52%  
Chemistry: Dye-primer ET; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.950319  
Consensus quality: 162138 bases at least Q40  
Consensus quality: 163048 bases at least Q30  
Consensus quality: 163441 bases at least Q20  
Insert size: 169000; agarose-fp  
Insert size: 165355; sum-of-contigs  
Quality coverage: 8.62 in Q20 bases; agarose-fp  
Quality coverage: 8.84 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: this is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

FEATURES  
source  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
BASE COUNT  
ORIGIN

1. 16967: contig of 16967 bp in length  
1. 16968: gap of unknown length  
1. 17068: contig of 31395 bp in length  
48463: gap of unknown length  
48563: contig of 50812 bp in length  
99375: gap of unknown length  
99474: gap of unknown length  
99475: 164281: contig of 64807 bp in length.  
Location/Qualifiers  
1. .164281  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/clone="RP11-603F24"  
1. .16967  
/note="assembly\_name:Contig98"  
17068. .48462  
/note="assembly\_name:Contig9"  
48563. .99374  
/note="assembly\_name:Contig10  
clone\_end:T7  
vector\_side:right"  
99475. .164281  
/note="assembly\_name:Contig11  
clone\_end:SP6  
vector\_side:left"  
46622 a 33513 c 33986 g 49860 t 300 others

Query Match 19.0%; Score 198; DB 2; Length 164281;  
Best Local Similarity 100.0%; Pred. No. 1.8e-98;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 455 ttccaggagatgcctacaccgattatgtacgtacgagatgtaccgagctctgaacttc 514  
Db 142462 TTCCAGGAGATGCCTACACCGATTATGTAGCTACGAGATGTACCGAGCTCTGAACCTC 142521  
QY 515 ttgtggagatactcaagtatggtcttcagtcgatatatggctattgtgtgttttg 574  
Db 142522 TTGTGGGAGATGACTCAGATATGGTCTTCAGTCGATATATGGCTATGTTGTTGTTT 142581  
QY 575 cagagctctctgacagcgcagccactgtgacctggaataacagatgtggaccaacttattc 634  
Db 142582 CAGAGCTCTCTGACAGCGCCAGCCACTGTGGCCTGGAAATCAGATGTGGACCACTTTATC 142641  
QY 635 tgataatcagaacactag 652  
Db 142642 TGATATCAGACACTAG 142659  
RESULT 13  
AC037489/c standard: DNA; HTG; 155028 BP.  
ID AC037489  
XX AC037489;  
AC AC037489;  
SV AC037489.2  
XX 10-APR-2000 (Rel. 63, Created)  
DT 14-MAY-2000 (Rel. 63, Last updated, Version 2)  
XX Homo sapiens chromosome 2 clone RP11-603F24 map 2, WORKING DRAFT SEQUENCE,  
DE 26 unordered pieces.  
XX HTGS\_DRAFT; HTGS\_PHASE1.  
XX Homo sapiens (human)  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
XX [1]  
RP 1-155028  
RA Birren B., Linton L., Nusbaum C., Lander E., Abraham H., Allen N.,  
RT "Homo sapiens chromosome 2, clone RP11-603F24";  
RL Unpublished.  
XX [2]  
RP 1-155028  
RA Birren B., Linton L., Nusbaum C., Lander E., Abraham H., Allen N.,  
RA Anderson S., Baldwin J., Barna N., Bastien V., Bada F., Boguslavskiy L.,  
RA Boukhaltier B., Brown A., Burkett G., Campopiano A., Castle A., Choepel Y.,  
RA Colangelo M., Collins S., Collymore A., Cooke P., DeArellano K., Dewar K.,  
RA Diaz J.S., Dodge S., Domino M., Doyle M., Ferrelira P., Fitzhugh W.,  
RA Gage D., Galagan J., Gardyna S., Ginde S., Goyette M., Graham L.,  
RA Grand-pierre N., Grant G., Hagos B., Heaford A., Horton L., Howland J.C.,  
RA Iliev I., Johnson R., Jones C., Kann L., Karatas A., Klein J., LaRoque K.,  
RA Lamazares R., Landers T., Lehoczy J., Levine R., Liu C., Liu G.,  
RA Locke K., Macdonald P., Marquis N., McCarthy M., McEwan P., McGurk A.,  
RA McKernan K., McPheeters R., Meldrim J., Meneus L., Mihova T., Miranda C.,  
RA Menga V., Morrow J., Murphy T., Maylor J., Norman C.H., O'Connor T.,  
RA O'Donnell P., O'Neil D., Olivar T.M., Oliver J., Peterson K., Pierre N.,  
RA Pisani C., Pollara V., Raymond C., Riley R., Rogov P., Rothman D., Roy A.,  
RA Santos R., Schauer S., Severy P., Spencer B., Stange-Thomann N.,  
RA Stojanovic N., Subramanian A., Talamas J., Tesfaye S., Theodore J.,  
RA Tirrell A., Travers M., Trigilio J., Vassiliev H., Viel R., Vo A.,  
RA Wilson B., Wu X., Wyman D., Ye W.J., Young G., Zainoun J., Zimmer A.,  
Zody M.;  
RT ; Submitted (09-APR-2000) to the EMBL/GenBank/DBJ databases.  
RL Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,  
RL Cambridge, MA 02141, USA  
XX

On May 12, 2000 this sequence version replaced gi:7528398.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www.seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: I9434  
 Center clone name: 603\_F\_24  
 ----- Summary Statistics  
 Sequencing vector: M13; M77815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 142960 bases at least Q40  
 Consensus quality: 148315 bases at least Q30  
 Consensus quality: 150779 bases at least Q20  
 Insert size: 164000; agarose-fp  
 Insert size: 152528; sum-of-contigs  
 Quality coverage: 4.4 in Q20 bases; agarose-fp  
 Quality coverage: 4.7 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 26 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 1104: contig of 1104 bp in length  
 \* 1105 1204: gap of 100 bp  
 \* 1205 2330: contig of 1126 bp in length  
 \* 2331 2430: gap of 100 bp  
 \* 2431 4965: contig of 2535 bp in length  
 \* 4966 5065: gap of 100 bp  
 \* 5066 7346: contig of 2281 bp in length  
 \* 7347 7446: gap of 100 bp  
 \* 7447 9679: contig of 2233 bp in length  
 \* 9680 9779: gap of 100 bp  
 \* 9780 11977: contig of 2198 bp in length  
 \* 11978 12077: gap of 100 bp  
 \* 12078 14038: contig of 1961 bp in length  
 \* 14039 14139: gap of 100 bp  
 \* 14139 15917: contig of 1779 bp in length  
 \* 15918 16017: gap of 100 bp  
 \* 16018 18776: contig of 2759 bp in length  
 \* 18777 18876: gap of 100 bp  
 \* 18877 21090: contig of 2214 bp in length  
 \* 21091 21190: gap of 100 bp  
 \* 21191 23923: contig of 2733 bp in length  
 \* 23924 24023: gap of 100 bp  
 \* 24024 27664: contig of 3641 bp in length  
 \* 27665 27764: gap of 100 bp  
 \* 27765 31535: contig of 3771 bp in length  
 \* 31536 31635: gap of 100 bp  
 \* 31636 36572: contig of 4937 bp in length  
 \* 36573 36672: gap of 100 bp  
 \* 36673 43155: contig of 6483 bp in length  
 \* 43156 43255: gap of 100 bp  
 \* 43256 48604: contig of 5349 bp in length  
 \* 48605 48704: gap of 100 bp  
 \* 48705 54578: contig of 5874 bp in length  
 \* 54579 54678: gap of 100 bp  
 \* 54679 61303: contig of 6625 bp in length  
 \* 61304 61403: gap of 100 bp  
 \* 61404 67988: contig of 6585 bp in length  
 \* 67989 68088: gap of 100 bp  
 \* 68089 75061: contig of 6973 bp in length  
 \* 75062 75161: gap of 100 bp  
 \* 75162 84322: contig of 9161 bp in length

CC \* 84323 84422: gap of 100 bp  
 CC \* 84423 93345: contig of 8923 bp in length  
 CC \* 93346 93445: gap of 100 bp  
 CC \* 93446 107415: contig of 13970 bp in length  
 CC \* 107416 107515: gap of 100 bp  
 CC \* 107516 119348: contig of 11833 bp in length  
 CC \* 119349 119448: gap of 100 bp  
 CC \* 119449 134591: contig of 15143 bp in length  
 CC \* 134592 134691: gap of 100 bp  
 CC \* 134692 155028: contig of 20337 bp in length.  
 XX  
 FH FH Location/Qualifiers  
 FT source  
 FT 1. .155028  
 FT /chromosome="2"  
 FT /db\_xref="taxon:9606"  
 FT /organism="Homo sapiens"  
 FT /map="2"  
 FT /clone\_lib="RPCI-11 Human Male BAC"  
 FT /clone="RP11-603F24"  
 FT 1. .1104  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 1205. .2330  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 2431. .4965  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 5066. .7346  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 7447. .9679  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 9780. .11977  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 12078. .14038  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 14139. .15917  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 16018. .18776  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 18877. .21090  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 21191. .23923  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 24024. .27664  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 27765. .31535  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 31636. .36572  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 36673. .43155  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 43256. .48604  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 48705. .54578  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 54679. .61303  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 61404. .67988  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 68089. .75061  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 75162. .84322  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 84423. .93345  
 FT /note="assembly\_fragment clone\_end:77 vector\_side:right"  
 FT misc\_feature  
 FT 93446. .107415  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 107516. .119348  
 FT /note="assembly\_fragment"  
 FT misc\_feature  
 FT 119449. .134591  
 FT /note="assembly\_fragment clone\_end:SP6 vector\_side:left"  
 FT misc\_feature  
 FT 134692. .155028  
 FT /note="assembly\_fragment"  
 FT XX

SQ Sequence 155028 BP; 43370 A; 31237 C; 31433 G; 46482 T; 2506 other;

Query Match 5.7%; Score 59; DB 33; Length 155028;  
Best Local Similarity 100.0%; Pred. No. 3.4e-21;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 831 ggaactcttgagaaagtctcagatgttcaatctgcttcgaacctcatgaagg 889  
||||| ggaactctttgagaaaagtctcagatgttcaatctgcttcgaacctcatgaagg 889  
Db 54111 GGAAACTCTTTGAGAAAAGTTCAGATGTTCATCTCTGGCTCTGAACCTCATGAAG 54053

RESULT 14  
GG0312524 10135 bp DNA PRI 13-FEB-2001  
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,  
DEFINITION isolate G227W.  
ACCESSION AJ312524  
VERSION AJ312524.1 GI:12830520  
KEYWORDS lowland gorilla.  
SOURCE Gorilla gorilla gorilla  
ORGANISM Gorilla gorilla gorilla  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Gorilla.

REFERENCE 1 (bases 1 to 10135)  
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.  
TITLE Great ape DNA sequences reveal a reduced diversity and an expansion  
JOURNAL Nat. Genet. 27 (2), 155-156 (2001)  
PUBMED 11175781  
AUTHORS Kaessmann,H.G.F.  
DIRECT SUBMISSION Submitted (09-FEB-2001) kaessmann H.G.F., Evolutionary Genetics  
(Prof. Paabo), Max-planck-Institute for Evolutionary Anthropology,  
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY

COMMENT non-coding sequence  
FEATURES Location/Qualifiers  
1..10135  
/organism="Gorilla gorilla gorilla"  
/isolate="WL287"  
/sub\_species="gorilla"  
/db\_xref="taxon:9595"  
/chromosome="X"  
/map="Xql3.3"  
/sex="male"

BASE COUNT 3233 a 2045 c 1802 g 3055 t  
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10135;  
Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagacggggtttgccattgttacaccagctggtctcgaaactc 696  
||||| tagtagacggggtttgccattgtttccaccagctggtctcgaaactc 9081  
Db 9035 TAGTAGACGGGGTTTCGCATTGTTCACCAGCTGGTCTCGAATC 9081

RESULT 16  
GG0312527 10135 bp DNA PRI 13-FEB-2001  
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,  
DEFINITION isolate G165.  
ACCESSION AJ312527  
VERSION AJ312527.1 GI:12830523  
KEYWORDS lowland gorilla.  
SOURCE Gorilla gorilla gorilla  
ORGANISM Gorilla gorilla gorilla  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Gorilla.

REFERENCE 1 (bases 1 to 10135)  
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.  
TITLE Great ape DNA sequences reveal a reduced diversity and an expansion  
JOURNAL Nat. Genet. 27 (2), 155-156 (2001)  
PUBMED 11175781  
AUTHORS Kaessmann,H.G.F.  
DIRECT SUBMISSION Submitted (09-FEB-2001) kaessmann H.G.F., Evolutionary Genetics  
(Prof. Paabo), Max-planck-Institute for Evolutionary Anthropology,  
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY

COMMENT non-coding sequence  
FEATURES Location/Qualifiers  
1..10135  
/organism="Gorilla gorilla gorilla"  
/isolate="G227W"  
/sub\_species="gorilla"  
/db\_xref="taxon:9595"  
/chromosome="X"  
/map="Xql3.3"  
/sex="male"

BASE COUNT 3233 a 2045 c 1802 g 3055 t  
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10135;  
Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagacggggtttgccattgttacaccagctggtctcgaaactc 696  
||||| tagtagacggggtttgccattgtttccaccagctggtctcgaaactc 9081  
Db 9035 TAGTAGACGGGGTTTCGCATTGTTCACCAGCTGGTCTCGAATC 9081

RESULT 15  
GG0312526 10135 bp DNA PRI 13-FEB-2001  
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,  
DEFINITION isolate WL287.  
ACCESSION AJ312526  
VERSION AJ312526.1 GI:12830522  
KEYWORDS lowland gorilla.  
SOURCE Gorilla gorilla gorilla  
ORGANISM Gorilla gorilla gorilla  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Gorilla.

REFERENCE 1 (bases 1 to 10135)  
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggtggtctcgaactc 696  
 |||  
 Db 9035 TAGTAGAGACGGGGTTTCGCCATGTTGACACGGCTGGTCTCGAACTC 9081

## RESULT 17

GG0312528 10135 bp DNA PRI 13-FEB-2001  
 LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,  
 DEFINITION isolate GI979.  
 ACCESSION AJ312528  
 VERSION AJ312528.1 GI:12830524  
 KEYWORDS lowland gorilla.  
 SOURCE Gorilla gorilla gorilla  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
 REFERENCE 1 (bases 1 to 10135)  
 AUTHORS Kaessmann, H., Wiebe, V., Weiss, G. and Paabo, S.  
 TITLE Great ape DNA sequences reveal a reduced diversity and an expansion  
 in humans  
 JOURNAL Nat. Genet. 27 (2), 155-156 (2001)  
 PUBMED 11175781  
 REFERENCE 2 (bases 1 to 10135)  
 AUTHORS Kaessmann, H. G. F.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-FEB-2001) Kaessmann H. G. F., Evolutionary Genetics  
 (Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,  
 Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY  
 COMMENT non-coding sequence.  
 FEATURES Location/Qualifiers  
 source  
 1. .10135  
 /organism="Gorilla gorilla gorilla"  
 /isolate="GI979"  
 /sub\_species="gorilla"  
 /db\_xref="taxon:9595"  
 /chromosome="X"  
 /map="Xql3.3"  
 /sex="male"

BASE COUNT 3233 a 2045 c 1802 g 3055 t  
 ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10135;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggtggtctcgaactc 696  
 |||  
 Db 9035 TAGTAGAGACGGGGTTTCGCCATGTTGACACGGCTGGTCTCGAACTC 9081

## RESULT 18

GG0241093 10139 bp DNA PRI 07-MAY-1999  
 LOCUS Gorilla gorilla Xql3.3 non-coding region, isolated from Western  
 DEFINITION lowland  
 ACCESSION AJ241093  
 VERSION AJ241093.1 GI:4775320  
 KEYWORDS non-coding region.  
 SOURCE gorilla.  
 ORGANISM Gorilla gorilla  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
 REFERENCE 1 (bases 1 to 10139)  
 AUTHORS Kaessmann, H., Helssig, F., von Haeseler, A. and Paabo, S.  
 TITLE DNA sequence variation in a non-coding region of low recombination  
 on the human X chromosome  
 JOURNAL Nat. Genet. 22 (1), 78-81 (1999)  
 MEDLINE 99251584  
 REFERENCE 2 (bases 1 to 10139)

AUTHORS Kaessmann, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-MAR-1999) Kaessmann H, Max-Planck-Institute of  
 Evolutionary Anthropology Inselstrasse 22, 04103 Leipzig, FRG  
 FEATURES Location/Qualifiers  
 source  
 1. .10139  
 /organism="Gorilla gorilla"  
 /isolate="from Western lowland"  
 /db\_xref="taxon:9593"  
 /chromosome="X"  
 /map="Xql3.3"  
 misc\_feature 1. .10139  
 /note="non-coding region"  
 BASE COUNT 3234 a 2044 c 1803 g 3058 t  
 ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10139;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggtggtctcgaactc 696  
 |||  
 Db 9038 TAGTAGAGACGGGGTTTCGCCATGTTGACACGGCTGGTCTCGAACTC 9084

## RESULT 19

GG0312522 10139 bp DNA PRI 13-FEB-2001  
 LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,  
 DEFINITION isolate GI07W.  
 ACCESSION AJ312522  
 VERSION AJ312522.1 GI:12830518  
 KEYWORDS lowland gorilla.  
 SOURCE Gorilla gorilla gorilla  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
 REFERENCE 1 (bases 1 to 10139)  
 AUTHORS Kaessmann, H., Wiebe, V., Weiss, G. and Paabo, S.  
 TITLE Great ape DNA sequences reveal a reduced diversity and an expansion  
 in humans  
 JOURNAL Nat. Genet. 27 (2), 155-156 (2001)  
 PUBMED 11175781  
 REFERENCE 2 (bases 1 to 10139)  
 AUTHORS Kaessmann, H. G. F.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-FEB-2001) Kaessmann H. G. F., Evolutionary Genetics  
 (Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,  
 Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY  
 COMMENT non-coding sequence.  
 FEATURES Location/Qualifiers  
 source  
 1. .10139  
 /organism="Gorilla gorilla gorilla"  
 /isolate="GI07W"  
 /sub\_species="gorilla"  
 /db\_xref="taxon:9595"  
 /chromosome="X"  
 /map="Xql3.3"  
 /sex="male"

BASE COUNT 3238 a 2043 c 1797 g 3061 t  
 ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10139;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggtggtctcgaactc 696  
 |||  
 Db 9039 TAGTAGAGACGGGGTTTCGCCATGTTGACACGGCTGGTCTCGAACTC 9085

```
RESULT 20
GGO312529 10139 bp DNA PRI 13-FEB-2001
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
DEFINITION isolate G014.
ACCESSION AJ312529
VERSION AJ312529.1 GI:12830525
SOURCE lowland gorilla.
ORGANISM Gorilla gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE 1 (bases 1 to 10139)
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE Great ape DNA sequences reveal a reduced diversity and an expansion
in humans
JOURNAL Nat. Genet. 27 (2), 155-156 (2001)
PUBMED 11175781
REFERENCE 2 (bases 1 to 10139)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
(Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT non-coding sequence.
FEATURES
source
1..10139
/organism="Gorilla gorilla gorilla"
/isolate="G014"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT 3238 a 2043 c 1797 g 3061 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10139;
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
DB 9039 TAGTAGAGACGGGTTTCGCCATGTTGACACAGGCTGCTCGAACTC 9085

RESULT 21
GGO312530 10139 bp DNA PRI 13-FEB-2001
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
DEFINITION isolate G1530.
ACCESSION AJ312530
VERSION AJ312530.1 GI:12830526
SOURCE lowland gorilla.
ORGANISM Gorilla gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE 1 (bases 1 to 10139)
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE Great ape DNA sequences reveal a reduced diversity and an expansion
in humans
JOURNAL Nat. Genet. 27 (2), 155-156 (2001)
PUBMED 11175781
REFERENCE 2 (bases 1 to 10139)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
(Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT non-coding sequence.
FEATURES
source
1..10139
/organism="Gorilla gorilla gorilla"
/isolate="G1530"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT 3234 a 2048 c 1802 g 3057 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10141;
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
DB 9041 TAGTAGAGACGGGTTTCGCCATGTTGACACAGGCTGCTCGAACTC 9087

RESULT 23
GGO312523 10142 bp DNA PRI 13-FEB-2001
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
DEFINITION isolate G167W.
COMMENT non-coding sequence.
FEATURES
source
1..10141
/organism="Gorilla gorilla gorilla"
/isolate="G1537W"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT 3238 a 2043 c 1797 g 3061 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10139;
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
DB 9039 TAGTAGAGACGGGTTTCGCCATGTTGACACAGGCTGCTCGAACTC 9085
```

```
1..10139
/organism="Gorilla gorilla gorilla"
/isolate="G1530"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT 3238 a 2043 c 1797 g 3061 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10139;
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
DB 9039 TAGTAGAGACGGGTTTCGCCATGTTGACACAGGCTGCTCGAACTC 9085

RESULT 22
GGO312521 10141 bp DNA PRI 13-FEB-2001
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
DEFINITION isolate G1537W.
ACCESSION AJ312521
VERSION AJ312521.1 GI:12830517
SOURCE lowland gorilla.
ORGANISM Gorilla gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE 1 (bases 1 to 10141)
AUTHORS Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE Great ape DNA sequences reveal a reduced diversity and an expansion
in humans
JOURNAL Nat. Genet. 27 (2), 155-156 (2001)
PUBMED 11175781
REFERENCE 2 (bases 1 to 10141)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
(Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
COMMENT non-coding sequence.
FEATURES
source
1..10141
/organism="Gorilla gorilla gorilla"
/isolate="G1537W"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT 3234 a 2048 c 1802 g 3057 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10141;
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
DB 9041 TAGTAGAGACGGGTTTCGCCATGTTGACACAGGCTGCTCGAACTC 9087

RESULT 23
GGO312523 10142 bp DNA PRI 13-FEB-2001
LOCUS Gorilla gorilla gorilla Xql3.3 chromosome non-coding sequence,
DEFINITION isolate G167W.
COMMENT non-coding sequence.
FEATURES
source
1..10141
/organism="Gorilla gorilla gorilla"
/isolate="G1537W"
/sub_species="gorilla"
/db_xref="taxon:9595"
/chromosome="X"
/map="Xql3.3"
/sex="male"
BASE COUNT 3234 a 2048 c 1802 g 3057 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10141;
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;
QY 650 tagtagagacgggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
DB 9041 TAGTAGAGACGGGTTTCGCCATGTTGACACAGGCTGCTCGAACTC 9087
```

```

ACCESSION      AJ312523
VERSION        AJ312523.1  GI:12830519
KEYWORDS
SOURCE         lowland gorilla.
ORGANISM       Gorilla gorilla gorilla
               Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Eukaryota; Eutheria; Primates; Catarrhini; Hominiidae; Gorilla.
REFERENCE
AUTHORS        Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE          Great ape DNA sequences reveal a reduced diversity and an expansion
               in humans
JOURNAL        Nat. Genet. 27 (2), 155-156 (2001)
PUBMED         11175781
REFERENCE      2 (bases 1 to 10142)
AUTHORS        Kaessmann,H.G.F.
TITLE          Direct Submission
JOURNAL        Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
               (Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
               Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
               non-coding sequence.
FEATURES       Location/Qualifiers
               1..10142
               /organism="Gorilla gorilla gorilla"
               /isolate="G167W"
               /sub_species="gorilla"
               /db_xref="taxon:9595"
               /chromosome="X"
               /map="Xq13.3"
               /sex="male"
BASE COUNT    3237 a 2044 c 1799 g 3062 t
ORIGIN
Query Match   4.5%; Score 47; DB 9; Length 10142;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagacggggttcgccatgtgaccagctggtctcgaactc 696
|||||
DB 9042 TAGTAGAGACGGGTTTCGCCATGTTGACCAGCGTGGTCTCGAATC 9088

RESULT 24
GG0312525
LOCUS          GG0312525 10142 bp DNA PRI 13-FEB-2001
DEFINITION     Gorilla gorilla beringei Xq13.3 chromosome non-coding sequence,
               Isolate Mou074.
ACCESSION      AJ312525
VERSION        AJ312525.1  GI:12830521
KEYWORDS
SOURCE         mountain gorilla.
ORGANISM       Gorilla gorilla beringei
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Gorilla.
REFERENCE
AUTHORS        Kaessmann,H., Wiebe,V., Weiss,G. and Paabo,S.
TITLE          Great ape DNA sequences reveal a reduced diversity and an expansion
               in humans
JOURNAL        Nat. Genet. 27 (2), 155-156 (2001)
PUBMED         11175781
REFERENCE      2 (bases 1 to 10142)
AUTHORS        Kaessmann,H.G.F.
TITLE          Direct Submission
JOURNAL        Submitted (09-FEB-2001) Kaessmann H.G.F., Evolutionary Genetics
               (Prof. Paabo), Max-Planck-Institute for Evolutionary Anthropology,
               Inselstrasse 22, Leipzig, Saxony, 04103, GERMANY
               non-coding sequence.
FEATURES       Location/Qualifiers
               1..10142
               /organism="Gorilla gorilla beringei"
               /isolate="Mou074"
               /sub_species="beringei"
               /db_xref="taxon:9594"

```

```

/chromosome="X"
/map="Xq13.3"
/sex="male"
BASE COUNT    3238 a 2044 c 1799 g 3061 t
ORIGIN
Query Match   4.5%; Score 47; DB 9; Length 10142;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagacggggttcgccatgtgaccagctggtctcgaactc 696
|||||
DB 9042 TAGTAGAGACGGGTTTCGCCATGTTGACCAGCGTGGTCTCGAATC 9088

RESULT 25
PPA270090
LOCUS          PPA270090 10145 bp DNA PRI 05-NOV-1999
DEFINITION     Pan paniscus Xq13.3 non-coding region, B1.
ACCESSION      AJ270090
VERSION        AJ270090.1  GI:6273615
KEYWORDS       non-coding region.
SOURCE         pygmy chimpanzee.
ORGANISM       Pan paniscus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
REFERENCE
AUTHORS        Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE          Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL        Science 286 (5442), 1159-1162 (1999)
MEDLINE        20018332
REFERENCE      2 (bases 1 to 10145)
AUTHORS        Kaessmann,H.G.F.
TITLE          Direct Submission
JOURNAL        Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
               Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
               Leipzig, 04103, GERMANY
FEATURES       Location/Qualifiers
               1..10145
               /organism="Pan paniscus"
               /db_xref="taxon:9597"
               /country="Democratic Republic of the Congo"
               /chromosome="X"
               /sex="male"
               /map="Xq13.3"
               /note="B1"
               misc_feature 1..10145
               /note="non-coding DNA"
BASE COUNT    3226 a 2031 c 1812 g 3076 t
ORIGIN
Query Match   4.5%; Score 47; DB 9; Length 10145;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagacggggttcgccatgtgaccagctggtctcgaactc 696
|||||
DB 9045 TAGTAGAGACGGGTTTCGCCATGTTGACCAGCGTGGTCTCGAATC 9091

RESULT 26
PPA270091
LOCUS          PPA270091 10145 bp DNA PRI 05-NOV-1999
DEFINITION     Pan paniscus Xq13.3 non-coding region, B4.
ACCESSION      AJ270091
VERSION        AJ270091.1  GI:6273616
KEYWORDS       non-coding region.
SOURCE         pygmy chimpanzee.
ORGANISM       Pan paniscus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.

```



us-09-671-050-9.oli20.rge

Fri May 3 11:02:09 2002

Best Local Similarity 100.0%; Pred. No. 1.5e-14; Mismatches 0; Indels 0; Gaps 0;  
Matches 47; Conservative 0;

QY 650 taatagacaggggttcgccatgttgaccagggtggtctcgaaactc 696  
|||||  
DB 9045 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9091

RESULT 28  
PPA270094 10145 bp DNA PRI 05-NOV-1999  
LOCUS Pan paniscus Xq13.3 non-coding region, B5.  
DEFINITION AJ270094  
ACCESSION AJ270094  
VERSION AJ270094.1 GI:6273638  
KEYWORDS non-coding region.  
SOURCE pygmy chimpanzee.  
ORGANISM Pan paniscus

REFERENCE  
AUTHORS Kaessmann, H., Wiebe, V. and Paabo, S.  
TITLE Extensive nuclear DNA sequence diversity among chimpanzees  
JOURNAL Science 286 (5442), 1159-1162 (1999)  
MEDLINE 20018332  
REFERENCE 2 (bases 1 to 10145)  
AUTHORS Kaessmann, H.G.F.  
TITLE Direct Submission  
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo, Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22, Leipzig, 04103, GERMANY

FEATURES  
source  
1..10145  
/organism="Pan paniscus"  
/db\_xref="taxon:9597"  
/country="Democratic Republic of the Congo"  
/chromosome="X"  
/sex="male"  
/map="Xq13.3"  
/note="B4"  
misc\_feature 1..10145  
/note="non-coding DNA"  
BASE COUNT 3224 a 2032 c 1812 g 3077 t  
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10145;  
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Mismatches 0; Indels 0; Gaps 0;  
Matches 47; Conservative 0;

QY 650 taatagacaggggttcgccatgttgaccagggtggtctcgaaactc 696  
|||||  
DB 9045 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9091

RESULT 27  
PPA270093 10145 bp DNA PRI 05-NOV-1999  
LOCUS Pan paniscus Xq13.3 non-coding region, B2.  
DEFINITION AJ270093  
ACCESSION AJ270093  
VERSION AJ270093.1 GI:6273618  
KEYWORDS non-coding region.  
SOURCE pygmy chimpanzee.  
ORGANISM Pan paniscus

REFERENCE  
AUTHORS Kaessmann, H., Wiebe, V. and Paabo, S.  
TITLE Extensive nuclear DNA sequence diversity among chimpanzees  
JOURNAL Science 286 (5442), 1159-1162 (1999)  
MEDLINE 20018332  
REFERENCE 2 (bases 1 to 10145)  
AUTHORS Kaessmann, H.G.F.  
TITLE Direct Submission  
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo, Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22, Leipzig, 04103, GERMANY

FEATURES  
source  
1..10145  
/organism="Pan paniscus"  
/db\_xref="taxon:9597"  
/country="Democratic Republic of the Congo"  
/chromosome="X"  
/sex="male"  
/map="Xq13.3"  
/note="B2"  
misc\_feature 1..10145  
/note="non-coding DNA"  
BASE COUNT 3224 a 2032 c 1812 g 3077 t  
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10145;  
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Mismatches 0; Indels 0; Gaps 0;  
Matches 47; Conservative 0;

QY 650 taatagacaggggttcgccatgttgaccagggtggtctcgaaactc 696  
|||||  
DB 9045 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9091

RESULT 29  
PPA270092 10151 bp DNA PRI 05-NOV-1999  
LOCUS Pan paniscus Xq13.3 non-coding region, B5.  
DEFINITION AJ270092  
ACCESSION AJ270092  
VERSION AJ270092.1 GI:6273617  
KEYWORDS non-coding region.  
SOURCE pygmy chimpanzee.  
ORGANISM Pan paniscus

REFERENCE  
AUTHORS Kaessmann, H., Wiebe, V. and Paabo, S.  
TITLE Extensive nuclear DNA sequence diversity among chimpanzees  
JOURNAL Science 286 (5442), 1159-1162 (1999)  
MEDLINE 20018332  
REFERENCE 2 (bases 1 to 10151)  
AUTHORS Kaessmann, H.G.F.  
TITLE Direct Submission

Query Match 4.5%; Score 47; DB 9; Length 10145;

JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,  
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,  
Leipzig, 04103, GERMANY

FEATURES  
source Location/Qualifiers

1..10151  
/organism="Pan paniscus"  
/db\_xref="taxon:9597"  
/country="Democratic Republic of the Congo"  
/chromosome="X"  
/sex="male"  
/map="Xq13.3"  
/note="Bs"  
misc\_feature 1..10151  
/note="non-coding DNA"  
BASE COUNT 3227 a 2032 c 1811 g 3081 t  
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151;  
Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagacgggtttcccatgttgaccaggctgtctgaactc 696  
|||||  
Db 9051 TAGTAGACGGGTTTCGCCATGTTGACCAGGCTGCTCGAACTC 9097

RESULT 30  
PTR270065  
LOCUS Pan troglodytes troglodytes Xq13.3 non-coding region, C2.  
DEFINITION Pan troglodytes troglodytes Xq13.3 non-coding region, C2.  
ACCESSION AJ270065  
VERSION AJ270065.1 GI:6273630  
KEYWORDS non-coding region.  
SOURCE chimpanzee.  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
AUTHORS Kaessmann, H., Wiebe, V. and Paabo, S.  
TITLE Extensive nuclear DNA sequence diversity among chimpanzees  
JOURNAL Science 286 (5442), 1159-1162 (1999)  
MEDLINE 20018332  
REFERENCE 2 (bases 1 to 10151)  
AUTHORS Kaessmann, H.G.F.  
TITLE Direct Submission  
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,  
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,  
Leipzig, 04103, GERMANY  
FEATURES  
source Location/Qualifiers

1..10151  
/organism="Pan troglodytes"  
/sub\_species="troglodytes"  
/db\_xref="taxon:9598"  
/country="Gabon"  
/chromosome="X"  
/sex="male"  
/map="Xq13.3"  
/note="C2"  
misc\_feature 1..10151  
/note="non-coding DNA"  
BASE COUNT 3221 a 2033 c 1814 g 3083 t  
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151;  
Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagacgggtttcccatgttgaccaggctgtctgaactc 696  
|||||  
Db 9051 TAGTAGACGGGTTTCGCCATGTTGACCAGGCTGCTCGAACTC 9097

RESULT 31  
PTR270066

LOCUS Pan troglodytes troglodytes Xq13.3 non-coding region, C3.  
DEFINITION Pan troglodytes troglodytes Xq13.3 non-coding region, C3.  
ACCESSION AJ270066  
VERSION AJ270066.1 GI:6273631  
KEYWORDS non-coding region.  
SOURCE chimpanzee.  
ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
AUTHORS Kaessmann, H., Wiebe, V. and Paabo, S.  
TITLE Extensive nuclear DNA sequence diversity among chimpanzees  
JOURNAL Science 286 (5442), 1159-1162 (1999)  
MEDLINE 20018332  
REFERENCE 2 (bases 1 to 10151)  
AUTHORS Kaessmann, H.G.F.  
TITLE Direct Submission  
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,  
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,  
Leipzig, 04103, GERMANY  
FEATURES  
source Location/Qualifiers

1..10151  
/organism="Pan troglodytes"  
/sub\_species="troglodytes"  
/db\_xref="taxon:9598"  
/country="Gabon"  
/chromosome="X"  
/sex="male"  
/map="Xq13.3"  
/note="C3"  
misc\_feature 1..10151  
/note="non-coding DNA"  
BASE COUNT 3221 a 2031 c 1817 g 3082 t  
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10151;  
Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 tagtagacgggtttcccatgttgaccaggctgtctgaactc 696  
|||||  
Db 9051 TAGTAGACGGGTTTCGCCATGTTGACCAGGCTGCTCGAACTC 9097

RESULT 32  
PTR270068  
LOCUS Pan troglodytes troglodytes Xq13.3 non-coding region, C7.  
DEFINITION Pan troglodytes troglodytes Xq13.3 non-coding region, C7.  
ACCESSION AJ270068  
VERSION AJ270068.1 GI:6273633  
KEYWORDS non-coding region.  
SOURCE chimpanzee.  
ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
AUTHORS Kaessmann, H., Wiebe, V. and Paabo, S.  
TITLE Extensive nuclear DNA sequence diversity among chimpanzees  
JOURNAL Science 286 (5442), 1159-1162 (1999)  
MEDLINE 20018332  
REFERENCE 2 (bases 1 to 10151)  
AUTHORS Kaessmann, H.G.F.  
TITLE Direct Submission  
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,  
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,  
Leipzig, 04103, GERMANY  
FEATURES  
source Location/Qualifiers

1..10151  
/organism="Pan troglodytes"

```

/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="C7"

misc_feature
1..10151
/note="non-coding DNA"

BASE COUNT      3225 a  2031 c  1813 g  3082 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10151;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccagcgtggtctcgaaactc 696
|||||
Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 33
PTR270070
LOCUS      PTR270070 10151 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes troglodytes Xq13.3 non-coding region, C4.
ACCESSION AJ270070
VERSION AJ270070.1 GI:6273635
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 10151)
Kaessmann, H., Wiebe, V. and Paabo, S.
Extensive nuclear DNA sequence diversity among chimpanzees
Science 286 (5442), 1159-1162 (1999)
20018332
REFERENCE
2 (bases 1 to 10151)
Kaessmann, H.G.F.
Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
Pan troglodytes
Location/Qualifiers
1..10151
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="C8"

misc_feature
1..10151
/note="non-coding DNA"

BASE COUNT      3221 a  2032 c  1818 g  3080 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10151;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccagcgtggtctcgaaactc 696
|||||
Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 35
PTR270072
LOCUS      PTR270072 10151 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes troglodytes Xq13.3 non-coding region, C6.
ACCESSION AJ270072
VERSION AJ270072.1 GI:6273637
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 10151)
Kaessmann, H., Wiebe, V. and Paabo, S.
Extensive nuclear DNA sequence diversity among chimpanzees
Science 286 (5442), 1159-1162 (1999)
20018332
REFERENCE
2 (bases 1 to 10151)
Kaessmann, H.G.F.
Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
Pan troglodytes
Location/Qualifiers
1..10151
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"

misc_feature
1..10151
/note="non-coding DNA"

BASE COUNT      3223 a  2031 c  1818 g  3079 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10151;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacggggttcgccatgttgaccagcgtggtctcgaaactc 696
|||||
Db 9051 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACTC 9097

RESULT 34
PTR270071
LOCUS      PTR270071 10151 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes troglodytes Xq13.3 non-coding region, C8.
```

```

misc_feature      /map="Xq13.3"
                  /note="C6"
BASE COUNT      3222 a 2032 c 1815 g 3082 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10151;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tactagacgggtttgcacatgttgaccaggctggtcgcgaactc 696
|||||
Db 9051 TAGTAGACGGGGTTTCGCCATGTGTGACCAGGCTGCTCGAAGTC 9097

RESULT 36
PTR241092
LOCUS
DEFINITION      PTR241092 10152 bp DNA PRI 07-MAY-1999
ACCESSION      AJ241092
VERSION      AJ241092.1 GI:4775559
KEYWORDS      non-coding region.
SOURCE      Pan troglodytes verus.
ORGANISM      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
AUTHORS      Kaessmann,H., Hellwig,F., von Haeseler,A. and Paabo,S.
TITLE      DNA sequence variation in a non-coding region of low recombination
JOURNAL      Nat. Genet. 22 (1), 78-81 (1999)
MEDLINE      99251584
AUTHORS      Kaessmann,H.
TITLE      Direct Submission
JOURNAL      Submitted (22-MAR-1999), Kaessmann H, Max-Planck-Institute of
Evolutionary Anthropology Inselstrasse 22, 04103 Leipzig, FRG
FEATURES
source      1..10152
/organism="Pan troglodytes verus"
/sub_species="verus"
/isolate="from West Africa"
/db_xref="taxon:37012"
/chromosome="X"
/map="Xq13.3"
misc_feature      1..10152
/organism="Pan troglodytes verus"
/notes="non-coding region"
BASE COUNT      3229 a 2028 c 1814 g 3081 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10152;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tactagacgggtttgcacatgttgaccaggctggtcgcgaactc 696
|||||
Db 9052 TAGTAGACGGGGTTTCGCCATGTGTGACCAGGCTGCTCGAAGTC 9098

RESULT 37
PTR270061
LOCUS
DEFINITION      PTR270061 10154 bp DNA PRI 12-FEB-2001
ACCESSION      AJ270061
VERSION      AJ270061.1 GI:6273626
KEYWORDS      non-coding region.
SOURCE      Pan troglodytes troglodytes.
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

```

```

REFERENCE
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
TITLE      1 (bases 1 to 10154)
JOURNAL      Kaessmann,H., Wiebe,V. and Paabo,S.
MEDLINE      Extensive nuclear DNA sequence diversity among chimpanzees
20018332
REFERENCE
AUTHORS      2 (bases 1 to 10154)
TITLE      Kaessmann,H.G.F.
JOURNAL      Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES
source      Location/Qualifiers
1..10154
/organism="Pan troglodytes troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:37011"
/chromosome="X"
/map="Xq13.3"
/sex="male"
/country="Gabon"
/note="C11"
1..10154
misc_feature      /note="non-coding DNA"
BASE COUNT      3228 a 2030 c 1815 g 3081 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagacgggtttgcacatgttgaccaggctggtcgcgaactc 696
|||||
Db 9054 TAGTAGACGGGGTTTCGCCATGTGTGACCAGGCTGCTCGAAGTC 9100

RESULT 38
PTR270062
LOCUS
DEFINITION      PTR270062 10154 bp DNA PRI 05-NOV-1999
ACCESSION      AJ270062
VERSION      AJ270062.1 GI:6273627
KEYWORDS      non-coding region.
SOURCE      chimpanzee.
ORGANISM      Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
AUTHORS      1 (bases 1 to 10154)
TITLE      Kaessmann,H., Wiebe,V. and Paabo,S.
JOURNAL      Extensive nuclear DNA sequence diversity among chimpanzees
20018332
REFERENCE
AUTHORS      2 (bases 1 to 10154)
TITLE      Kaessmann,H.G.F.
JOURNAL      Direct Submission
Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES
source      Location/Qualifiers
1..10154
/organism="Pan troglodytes"
/sub_species="troglodytes"
/db_xref="taxon:9598"
/country="Gabon"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="C12"
1..10154
misc_feature      /note="non-coding DNA"
BASE COUNT      3227 a 2028 c 1814 g 3085 t
ORIGIN

```

us-09-671-050-9.oli20.rge

Fri May 3 11:02:09 2002

```

MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann,H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES
    source
        Location/Qualifiers
            1..10154
                /organism="Pan troglodytes"
                /sub_species="troglodytes"
                /db_xref="taxon:9598"
                /country="Gabon"
                /chromosome="X"
                /sex="male"
                /map="Xq13.3"
                /note="C1"
            1..10154
                /note="non-coding DNA"
        misc_feature 1..10154
BASE COUNT 3227 a 2029 c 1815 g 3083 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;

QY 650 tagtagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGCTCGAACTC 9100

RESULT 41
PTR270073 10154 bp DNA PRI 09-NOV-1999
LOCUS Pan troglodytes verus Xq13.3 non-coding region, w13.
DEFINITION AJ270073
ACCESSION AJ270073.1 GI:6318234
VERSION non-coding region.
KEYWORDS chimpanzee
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10154)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
Kaessmann,H.G.F.
Direct Submission
Submitted (27-Sep-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES
    source
        Location/Qualifiers
            1..10154
                /organism="Pan troglodytes"
                /sub_species="troglodytes"
                /db_xref="taxon:9598"
                /country="Gabon"
                /chromosome="X"
                /sex="male"
                /map="Xq13.3"
                /note="C10"
            1..10154
                /note="non-coding DNA"
        misc_feature 1..10154
BASE COUNT 3228 a 2030 c 1814 g 3082 t
ORIGIN
Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;

QY 650 tagtagacggggttcgccatgttgaccaggctggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAGGCTGCTCGAACTC 9100

RESULT 40
PTR270067 10154 bp DNA PRI 05-NOV-1999
LOCUS Pan troglodytes troglodytes Xq13.3 non-coding region, C1.
DEFINITION Pan troglodytes troglodytes
ACCESSION AJ270067
VERSION AJ270067.1 GI:6273632
KEYWORDS non-coding region.
SOURCE chimpanzee
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10154)
AUTHORS Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)

```

```

QY 650 tagtagacggggttcgccatgttgaccaggtggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTC 9100

RESULT 42
PTR270076
LOCUS PTR270076 10154 bp DNA PRI 08-NOV-1999
DEFINITION Pan troglodytes versus Xq13.3 non-coding region, w15.
ACCESSION AJ270076
VERSION AJ270076.1 GI:6318237
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
AUTHORS Kaessmann, H., Wiebe, V. and Paabo, S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann, H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES
source
1..10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="w15"
misc_feature
1..10154
/note="non-coding DNA"
BASE COUNT 3227 a 2027 c 1815 g 3085 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagacggggttcgccatgttgaccaggtggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTC 9100

RESULT 43
PTR270077
LOCUS PTR270077 10154 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes versus Xq13.3 non-coding region, w5.
ACCESSION AJ270077
VERSION AJ270077.1 GI:6273700
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
AUTHORS Kaessmann, H., Wiebe, V. and Paabo, S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann, H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY

```

```

FEATURES
source
1..10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="w5"
misc_feature
1..10154
/note="non-coding DNA"
BASE COUNT 3229 a 2026 c 1814 g 3085 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagacggggttcgccatgttgaccaggtggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTC 9100

RESULT 44
PTR270078
LOCUS PTR270078 10154 bp DNA PRI 05-NOV-1999
DEFINITION Pan troglodytes versus Xq13.3 non-coding region, w14.
ACCESSION AJ270078
VERSION AJ270078.1 GI:6273701
KEYWORDS non-coding region.
SOURCE chimpanzee.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
AUTHORS Kaessmann, H., Wiebe, V. and Paabo, S.
TITLE Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL Science 286 (5442), 1159-1162 (1999)
MEDLINE 20018332
REFERENCE 2 (bases 1 to 10154)
AUTHORS Kaessmann, H.G.F.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
Leipzig, 04103, GERMANY
FEATURES
source
1..10154
/organism="Pan troglodytes"
/sub_species="verus"
/db_xref="taxon:9598"
/country="Sierra Leone"
/chromosome="X"
/sex="male"
/map="Xq13.3"
/note="w14"
misc_feature
1..10154
/note="non-coding DNA"
BASE COUNT 3227 a 2028 c 1815 g 3084 t
ORIGIN

Query Match 4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagacggggttcgccatgttgaccaggtggtctcgaaactc 696
|||||
Db 9054 TAGTAGAGACGGGGTTTCGCCATGTTGACCAAGGCTGGTCTCGAACTC 9100

RESULT 45

```

Fri May 3 11:02:09 2002

/country="Sierra Leone"  
 /chromosome="X"  
 /sex="male"  
 /map="Xq13.3"  
 /note="W12"  
 misc\_feature 1..10154  
 /note="non-coding DNA"  
 BASE COUNT 3223 a 2030 c 1816 g 3085 t  
 ORIGIN  
 Query Match 4.5%; Score 47; DB 9; Length 10154;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 650 tagtagagacgggtttccgcatgttgaccaggctggtctcgaaactc 696  
 |||||  
 Db 9054 TAGTAGAGACGGGTTTCGCCATGTTGACACGAGCTGGTCTCGAATC 9100  
 |||||  
 RESULT 47  
 PTR270081 10154 bp DNA PRI 05-NOV-1999  
 LOCUS Pan troglodytes versus Xq13.3 non-coding region, w10.  
 DEFINITION  
 ACCESSION AJ270081  
 VERSION AJ270081.1 GI:6273704  
 KEYWORDS non-coding region.  
 SOURCE chimpanzee.  
 ORGANISM  
 Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 Kaessmann, H., Wiebe, V. and Paabo, S.  
 Extensive nuclear DNA sequence diversity among chimpanzees  
 Science 286 (5442), 1159-1162 (1999)  
 REFERENCE 20018332  
 MEDLINE  
 TITLE  
 AUTHORS  
 JOURNAL  
 REFERENCE 2 (bases 1 to 10154)  
 Kaessmann, H., Wiebe, V. and Paabo, S.  
 Extensive nuclear DNA sequence diversity among chimpanzees  
 Science 286 (5442), 1159-1162 (1999)  
 TITLE  
 AUTHORS  
 JOURNAL  
 FEATURES  
 source  
 1..10154  
 /organism="Pan troglodytes"  
 /sub\_species="verus"  
 /db\_xref="taxon:9598"  
 /country="Sierra Leone"  
 /chromosome="X"  
 /sex="male"  
 /map="Xq13.3"  
 /note="W6"  
 misc\_feature 1..10154  
 /note="non-coding DNA"  
 BASE COUNT 3228 a 2028 c 1814 g 3084 t  
 ORIGIN  
 Query Match 4.5%; Score 47; DB 9; Length 10154;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 650 tagtagagacgggtttccgcatgttgaccaggctggtctcgaaactc 696  
 |||||  
 Db 9054 TAGTAGAGACGGGTTTCGCCATGTTGACACGAGCTGGTCTCGAATC 9100  
 |||||  
 RESULT 46  
 PTR270080 10154 bp DNA PRI 05-NOV-1999  
 LOCUS Pan troglodytes versus Xq13.3 non-coding region, w12.  
 DEFINITION  
 ACCESSION AJ270080  
 VERSION AJ270080.1 GI:6273703  
 KEYWORDS non-coding region.  
 SOURCE chimpanzee.  
 ORGANISM  
 Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 Kaessmann, H., Wiebe, V. and Paabo, S.  
 Extensive nuclear DNA sequence diversity among chimpanzees  
 Science 286 (5442), 1159-1162 (1999)  
 REFERENCE 20018332  
 MEDLINE  
 TITLE  
 AUTHORS  
 JOURNAL  
 REFERENCE 2 (bases 1 to 10154)  
 Kaessmann, H.G.F.  
 Direct Submission  
 Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,  
 Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,  
 Leipzig, 04103, GERMANY  
 Location/Qualifiers  
 1..10154  
 /organism="Pan troglodytes"  
 /sub\_species="verus"  
 /db\_xref="taxon:9598"  
 /country="Sierra Leone"  
 /chromosome="X"  
 /sex="male"  
 /map="Xq13.3"  
 /note="W10"  
 misc\_feature 1..10154  
 /note="non-coding DNA"  
 BASE COUNT 3229 a 2028 c 1814 g 3083 t  
 ORIGIN  
 Query Match 4.5%; Score 47; DB 9; Length 10154;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 650 tagtagagacgggtttccgcatgttgaccaggctggtctcgaaactc 696  
 |||||  
 Db 9054 TAGTAGAGACGGGTTTCGCCATGTTGACACGAGCTGGTCTCGAATC 9100  
 |||||  
 RESULT 48  
 PTR270082 10154 bp DNA PRI 05-NOV-1999  
 LOCUS Pan troglodytes versus Xq13.3 non-coding region, w1.  
 DEFINITION  
 ACCESSION AJ270082  
 VERSION AJ270082.1 GI:6273705

/country="Sierra Leone"  
 /chromosome="X"  
 /sex="male"  
 /map="Xq13.3"  
 /note="W12"  
 misc\_feature 1..10154  
 /note="non-coding DNA"  
 BASE COUNT 3223 a 2030 c 1816 g 3085 t  
 ORIGIN  
 Query Match 4.5%; Score 47; DB 9; Length 10154;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 650 tagtagagacgggtttccgcatgttgaccaggctggtctcgaaactc 696  
 |||||  
 Db 9054 TAGTAGAGACGGGTTTCGCCATGTTGACACGAGCTGGTCTCGAATC 9100  
 |||||  
 RESULT 47  
 PTR270081 10154 bp DNA PRI 05-NOV-1999  
 LOCUS Pan troglodytes versus Xq13.3 non-coding region, w10.  
 DEFINITION  
 ACCESSION AJ270081  
 VERSION AJ270081.1 GI:6273704  
 KEYWORDS non-coding region.  
 SOURCE chimpanzee.  
 ORGANISM  
 Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 Kaessmann, H., Wiebe, V. and Paabo, S.  
 Extensive nuclear DNA sequence diversity among chimpanzees  
 Science 286 (5442), 1159-1162 (1999)  
 REFERENCE 20018332  
 MEDLINE  
 TITLE  
 AUTHORS  
 JOURNAL  
 REFERENCE 2 (bases 1 to 10154)  
 Kaessmann, H.G.F.  
 Direct Submission  
 Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,  
 Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,  
 Leipzig, 04103, GERMANY  
 Location/Qualifiers  
 1..10154  
 /organism="Pan troglodytes"  
 /sub\_species="verus"  
 /db\_xref="taxon:9598"  
 /country="Sierra Leone"  
 /chromosome="X"  
 /sex="male"  
 /map="Xq13.3"  
 /note="W10"  
 misc\_feature 1..10154  
 /note="non-coding DNA"  
 BASE COUNT 3229 a 2028 c 1814 g 3083 t  
 ORIGIN  
 Query Match 4.5%; Score 47; DB 9; Length 10154;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 650 tagtagagacgggtttccgcatgttgaccaggctggtctcgaaactc 696  
 |||||  
 Db 9054 TAGTAGAGACGGGTTTCGCCATGTTGACACGAGCTGGTCTCGAATC 9100  
 |||||  
 RESULT 48  
 PTR270082 10154 bp DNA PRI 05-NOV-1999  
 LOCUS Pan troglodytes versus Xq13.3 non-coding region, w1.  
 DEFINITION  
 ACCESSION AJ270082  
 VERSION AJ270082.1 GI:6273705

```

KEYWORDS    non-coding region.
SOURCE       chimpanzee.
ORGANISM     Pan troglodytes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE    1 (bases 1 to 10154)
AUTHORS      Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE        Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL      Science 286 (5442), 1159-1162 (1999)
MEDLINE      20018332
REFERENCE    2 (bases 1 to 10154)
AUTHORS      Kaessmann,H.G.F.
TITLE        Direct Submission
JOURNAL      Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
              Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
              Leipzig, 04103, GERMANY
FEATURES     source
              1..10154
              Location/Qualifiers
                /organism="Pan troglodytes"
                /sub_species="verus"
                /db_xref="taxon:9598"
                /country="Sierra Leone"
                /chromosome="X"
                /sex="male"
                /map="Xq13.3"
                /note="W1"
              1..10154
              /note="non-coding DNA"
BASE COUNT   3227 a 2030 c 1815 g 3082 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggtggtctcgaaactc 696
      |||||||
DB 9054 TAGTAGAGACGGGGTTTCCCATGTTGACCAAGCTGGTCTCGAACTC 9100

RESULT 49
PTR270083      10154 bp      DNA      PRI      05-NOV-1999
LOCUS          Pan troglodytes verus Xq13.3 non-coding region, w8.
DEFINITION     AJ270083
ACCESSION      AJ270083
VERSION        AJ270083.1 GI:6273706
KEYWORDS       non-coding region.
SOURCE         chimpanzee.
ORGANISM       Pan troglodytes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE      1 (bases 1 to 10154)
AUTHORS        Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE          Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL        Science 286 (5442), 1159-1162 (1999)
MEDLINE        20018332
REFERENCE      2 (bases 1 to 10154)
AUTHORS        Kaessmann,H.G.F.
TITLE          Direct Submission
JOURNAL        Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
              Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
              Leipzig, 04103, GERMANY
FEATURES       source
              1..10154
              Location/Qualifiers
                /organism="Pan troglodytes"
                /sub_species="verus"
                /db_xref="taxon:9598"
                /country="Sierra Leone"
                /chromosome="X"
                /sex="male"
                /map="Xq13.3"
                /note="W8"

```

```

misc_feature 1..10154
              /note="non-coding DNA"
BASE COUNT   3229 a 2028 c 1814 g 3083 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggtggtctcgaaactc 696
      |||||||
DB 9054 TAGTAGAGACGGGGTTTCCCATGTTGACCAAGCTGGTCTCGAACTC 9100

RESULT 50
PTR270084      10154 bp      DNA      PRI      05-NOV-1999
LOCUS          Pan troglodytes verus Xq13.3 non-coding region, w2.
DEFINITION     AJ270084
ACCESSION      AJ270084
VERSION        AJ270084.1 GI:6273707
KEYWORDS       non-coding region.
SOURCE         chimpanzee.
ORGANISM       Pan troglodytes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE      1 (bases 1 to 10154)
AUTHORS        Kaessmann,H., Wiebe,V. and Paabo,S.
TITLE          Extensive nuclear DNA sequence diversity among chimpanzees
JOURNAL        Science 286 (5442), 1159-1162 (1999)
MEDLINE        20018332
REFERENCE      2 (bases 1 to 10154)
AUTHORS        Kaessmann,H.G.F.
TITLE          Direct Submission
JOURNAL        Submitted (27-SEP-1999) Kaessmann H.G.F., Prof. Paabo,
              Max-Planck-Institute for Evolutionary Anthropology, Inselstr. 22,
              Leipzig, 04103, GERMANY
FEATURES       source
              1..10154
              Location/Qualifiers
                /organism="Pan troglodytes"
                /sub_species="verus"
                /db_xref="taxon:9598"
                /country="Sierra Leone"
                /chromosome="X"
                /sex="male"
                /map="Xq13.3"
                /note="W2"
              1..10154
              /note="non-coding DNA"
BASE COUNT   3228 a 2029 c 1815 g 3082 t
ORIGIN

Query Match      4.5%; Score 47; DB 9; Length 10154;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 tagtagagacgggttcgccatgttgaccaggtggtctcgaaactc 696
      |||||||
DB 9054 TAGTAGAGACGGGGTTTCCCATGTTGACCAAGCTGGTCTCGAACTC 9100

Search completed: May 2, 2002, 19:45:59
Job time: 9353 sec

```



us-09-671-050-9.oli20.rge

Fri May 3 11:02:09 2002



us-09-671-050-10.rag

Fri May 3 11:22:24 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:01:25 ; Search time 40.16 Seconds  
(without alignments)  
640.026 Million cell updates/sec

Title: US-09-671-050-10  
Perfect score: 1820  
Sequence: 1 MEKYEKLAKTGGSGYGVFK.....RKARNEGRNRQRQVLPLKS 347

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues 522463  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1820	100.0	347	22	AAE00494 Human kinase #5.
2	1796	98.7	356	22	AAE00491 Human kinase #2.
3	1636	89.9	315	22	AAE00495 Human kinase #6.
4	1612	88.6	324	22	AAE00492 Human kinase #3.
5	1593	87.5	360	22	AAU03525 Human protein kina
6	1270.5	69.8	296	22	AAU03525 Novel protein kina
7	1123	61.7	247	22	AAE05643 Novel protein kina
8	941.5	51.7	566	21	AAE05642 Rabbit KKIAMRE kin
9	791	43.5	187	22	AAE00490 Human kinase #4.
10	791	43.5	198	22	AAE00493 Human kinase #1.
11	758.5	41.7	455	21	AAU03525 Human lost in leuk

Human protein kina  
Rat lost in leukae  
Rat lost in leukae  
Novel protein kina  
Human polypeptide  
CDK2-green fluore  
Human polypeptide  
Homo sapiens Cdc2  
Green fluorescent  
Oryza sativa Cdc2  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Human protein kina  
Zea mays protein f  
Zea mays protein f  
Homo sapiens Cdc2  
Rattus norvegicus  
Cyclin-dependent k  
Pneumocystis carin  
Protein involved i  
Tomato cyclin-depe  
Cyclin dependent p  
Schizosaccharomyce  
Fission yeast prot  
Arabidopsis thalia  
Emeritella nidulan  
CDK1 protein. Can  
Candida albicans C  
Candida cyclin dep  
Ajellomyces capsul  
Arabidopsis thalia  
Arabidopsis thalia  
Human cyclin depen

ALIGNMENTS

RESULT 1  
AAE00494  
ID AAE00494 standard; Protein; 347 AA.

AC AAE00494;  
XX  
DT 19-JUN-2001 (first entry)  
XX  
DE Human kinase #5.  
XX  
XX Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder.  
XX  
OS Homo sapiens.

XX WO200123579-A1.  
XX 05-APR-2001.  
XX 27-SEP-2000; 2000WO-US26621.  
XX 28-SEP-1999; 99US-0156511.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX WPI: 2001-266166/27.  
XX N-PSDB; AAD03816.  
XX New isolated human kinase polynucleotide useful for generating  
XX antibodies, as reagents in diagnostic assays and for screening for  
XX compounds useful for treating mental, biological or medical diseases



|||||  
301 mnpddrltcsqlessyfsdfqeaqikrkrarnegnrrrqgv 342

RESULT 3  
AAE00495  
ID AAE00495 standard; Protein; 315 AA.

AC AAE00495;  
XX 19-JUN-2001 (first entry)  
DT Human kinase #6.  
DE Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder.

XX Homo sapiens.  
XX WO200123579-A1.  
XX 05-APR-2001.  
XX 27-SEP-2000; 2000WO-US26621.  
XX 28-SEP-1999; 99US-0156511.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI: 2001-266166/27.  
XX N-PSDB; AAD03817.  
XX New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases  
XX Claim 3; Page 33-34; 38pp; English.

XX The present sequence is novel human protein  
CC (NHP) known as human kinase. The human kinases share structural  
CC similarity with animal kinases, more particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.

XX Sequence 315 AA;

Query Match 89.9%; Score 1636; DB 22; Length 315;  
Best Local Similarity 90.8%; Pred. No. 2e-167;  
Matches 315; Conservative 0; Mismatches 0; Indels 32; Gaps 1;  
1 MEKYELAKTGEYSYGVVFKCRNKTSGQVAVKKFVESEDDPVVKKIALREIRMLKQKH 60  
|||||

Db 1 mekyelaktgeysygvvfkcrnktsgqvavkkfveseddpvvkkialreirmlkqlkh 60  
Qy 61 PNLVNLIEVFRKRKRKMLVFEYCDHTLLNELERNPNGVADGVITKSVLQOTLOALNFCIH 120  
Db 61 pnlvnlievfrkrkrkmlvfeycdhtllnelernpnngvadgviksvlwtlqalnfchih 120  
Qy 121 NCTHRDIPENILITKOGIICDFGAOILIPGDAYTDYVATRWYRAPELLVGDYQYS 180  
Db 121 ncthrdikpenililckgikicdfgaqilipgdaytdyvatrwyrappellvgdtqys 180  
Qy 181 SVDIWAIGCVFAELLTGQPLWPKGSVDQLYLILPTLVETGFRHVDQAGLELLTSSDPPA 240  
Db 181 svdiwaigcvfaelltgqplwpgksvdqlylilptlvetgfrhvdqagllelltssdppa 240  
Qy 241 VASOSAGITGKLIPIHQSIKSNFGFHGISIPEPEDMETLEEKFSVDHPVALNFMKGCCLK 300  
Db 218 -----gkllprhqsifkngfthgisipepedmetleekfsdvhvvalnfmkgcl 268  
Qy 301 MNPDDRITCSQLLESSYFDSFQEAQIKRKRARNEGRNRROOVLPLKS 347  
Db 269 mnpddrltcsqlessyfsdfqeaqikrkrarnegnrrrrqgvlpks 315

RESULT 4

AAE00492  
ID AAE00492 standard; Protein; 324 AA.

AC AAE00492;  
XX 19-JUN-2001 (first entry)  
DT Human kinase #3.  
DE Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder.  
XX Homo sapiens.  
XX WO200123579-A1.  
XX 05-APR-2001.  
XX 27-SEP-2000; 2000WO-US26621.  
XX 28-SEP-1999; 99US-0156511.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI: 2001-266166/27.  
XX N-PSDB; AAD03814.

XX New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases  
XX Disclosure; Page 30; 38pp; English.

XX The present sequence is novel human protein  
CC (NHP) known as human kinase. The human kinases share structural  
CC similarity with animal kinases, more particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes



immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;  
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;  
KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.  
XX  
OS Mus musculus.  
XX  
XX WO200073469-A2.  
XX  
XX 07-DEC-2000.  
XX  
XX 26-MAY-2000; 2000WO-US14842.  
XX  
XX 28-MAY-1999; 99US-0136503.  
XX  
XX (SUGE-) SUGEN INC.  
XX  
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;  
XX  
XX WPI; 2001-032161/04.  
XX  
XX N-PSDB; AAF44670.  
XX  
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers -  
XX  
XX Claim 10; Fig 1; 310pp; English.  
XX  
XX The present sequence is a novel protein kinase. The novel protein kinases  
CC and the nucleic acids that encode them may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC expression and activity. Diseases related to kinase expression and  
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related diseases, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
CC reproductive disorders.  
XX  
XX Sequence 296 AA;  
SQ  
Query Match 69.8%; Score 1270.5; DB 22; Length 296;  
Best Local Similarity 80.5%; Pred. NO. 4.1e-128;  
Matches 243; Conservative 19; Mismatches 5; Indels 35; Gaps 3;  
QY 46 KIALEIRMLKQKHPLNVLIEVFRKRKMLVPEYCDHTLLNELRNPNGVADGVKS 105  
DB 1 KIALEIRMLK-LKPLNVLNIEVFRKRKMLVFEYCDHTLLNELRNPNGVSGVKS 59  
QY 106 VLWQTLQALNFCNHNCIHRDIPENILITKQGIKICDFGAGIILPGDAYTDYVATRW 165  
DB 60 VLWQTLQALNFCNHNCIHRDIPENILITKQGIKICDFGAGIILPGDAYTDYVATRW 119  
QY 166 YRAPPELLVGTQYGVSSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLITRTIVETGFRHV 225  
DB 120 YRAPPELLVGTQYGVSSVDVWAGVGVFAELLTGQPLWPGKSDVDQLYLITRTIVETGFRHV 171  
QY 226 DOAGLELITSDPPAVASQSGAGITGKLIIPHQSIKSNCFHFGISIPEDMETLEKFS 285  
DB 172 -----gkllprhgsifrsngifrgisipepedmetleekfs 207  
QY 286 DVHPVALNFMKGCLKMNPDRLTCSQLLESSYFDSFQEAQIKRKARNEGRRRQ--QVL 343  
DB 208 NVQPVLSFMKGCLKMNPDRLTCAQLLSAYFESFQEDQMKRSEGRSRRRQNL 267

QY 344 PL 345  
DB 268 pl 269  
RESULT 7  
AAB65642  
ID AAB65642 standard; Protein; 247 AA.  
XX  
AC AAB65642;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE Novel protein kinase, SEQ ID NO: 169.  
XX  
KW Human; mouse; protein kinase; antiarthritic; antiarteritic; antisclerotic; osteopathic;  
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;  
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;  
KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.  
XX  
OS Homo sapiens.  
XX  
XX WO200073469-A2.  
XX  
XX 07-DEC-2000.  
XX  
XX 26-MAY-2000; 2000WO-US14842.  
XX  
XX 28-MAY-1999; 99US-0136503.  
XX  
XX (SUGE-) SUGEN INC.  
XX  
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;  
XX  
XX WPI; 2001-032161/04.  
XX  
XX N-PSDB; AAF44669.  
XX  
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers -  
XX  
XX Claim 10; Fig 1; 310pp; English.  
XX  
XX The present sequence is a novel protein kinase. The novel protein kinases  
CC and the nucleic acids that encode them may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC expression and activity. Diseases related to kinase expression and  
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related diseases, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
CC reproductive disorders.  
XX  
XX Sequence 247 AA;  
SQ  
Query Match 61.7%; Score 1123; DB 22; Length 247;  
Best Local Similarity 86.1%; Pred. NO. 2.3e-112;  
Matches 217; Conservative 1; Mismatches 0; Indels 34; Gaps 2;  
QY 96 NCVADGVKSVLWQTLQALNFCNHNCIHRDIPENILITKQGIKICDFGAGIILPGD 155  
DB 111

Db 1 ngvadgviqsvlwqtlqalnfnchihmcihrdikpenilitkqgikicdfgaqilpgd 60  
QY 156 AYTDYVATRWYRAPELLVGDFOYSSVDIWAIGCVFAELLTGTGPLWPKSDVDQLYLIIR 215  
Db 61 aytdyvatrwyrapelldvgtqygssvdiwaigcvfaelltgcplwpgksdvqlyliir 120  
QY 216 TLVETGFRHVDQAGLELLTSSDPPAVASOSAGITGKLIPIHQSIKSNFGFHGISIPEPE 275  
Db 121 t-----gklprhgsifksgffhgisipepe 148  
QY 276 DMETLEEKESDVHPVALNFMKGLKNPDDRLTCSQLESSEYDFEQEAQIKRKARNEGR 335  
Db 149 dmetleektsdvhpvalnfmkglknpddrltcsqlesseysdfseaqikrkarnegr 208  
QY 336 NRRRQ--QVLPL 345  
Db 209 nrrrqngqlpl 220  
RESULT 8  
ID AAY90724 standard; Protein; 566 AA.  
XX AC AAY90724;  
XX DT 15-AUG-2000 (first entry)  
XX DE Rabbit KTIAMRE kinase SEQ ID NO:4.  
XX KW Rabbit; KTIAMRE kinase; learning-induced kinase; learning; memory;  
KW cdc2-related kinase; brain; gene therapy; genetic disorder; detection;  
XX KW identification.  
XX OS Oryctolagus cuniculus.  
XX PN WO200020567-A2.  
XX PD 13-APR-2000.  
XX PF 01-OCT-1999; 99WO-US23010.  
XX PR 02-OCT-1998; 98US-0102906.  
XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.  
XX PI Thompson RF, Goni H, Sun W;  
XX DR WPI; 2000-328932/28.  
XX DR N-PSDB; AAA29745.  
XX PS Novel learning induced kinase polynucleotides and polypeptides, useful  
XX for the analysis of learning and memory, and for gene therapy -  
XX Claim 1; Fig 4; 64pp; English.  
XX CC The present sequence represents a learning-induced kinase, designated  
XX KTIAMRE kinase, which is isolated from rabbit brain tissue. KTIAMRE  
XX kinase is a cdc2-related kinase. The KTIAMRE kinase polynucleotides can  
XX be used to express recombinant protein for analysis, characterisation or  
XX therapeutic use, as markers for tissues in which the protein is  
XX preferentially expressed, as molecular weight markers on Southern gels,  
XX as chromosome markers or tags, to compare endogenous DNA sequences in  
XX patients to identify potential genetic disorders, as probes to hybridise  
XX and discover novel related sequences, as a source of PCR primers, and as  
XX an antigen to induce anti-DNA antibodies. The polypeptides can be used  
XX in assay to discover biological activity, to raise antibodies, as tissue  
XX markers, and to isolate correlative receptors or ligands. The  
XX polynucleotides may also be used for gene therapy for the treatment of  
XX disorders which are mediated by KTIAMRE kinase.  
XX SQ Sequence 566 AA;

Query Match 51.7%; Score 941.5; DB 21; Length 566;  
Best Local Similarity 49.7%; Pred. No. 3e-92;  
Matches 177; Conservative 67; Mismatches 65; Indels 47; Gaps 6;  
QY 1 MEKYEKLAKTGGSGYGVFKCRNKTSGQVAVKFFVESEDDPVVKIALREIRMLKOLKH 60  
Db 1 mekyenlgivgsgymvmckrknksgrivaikkkflesdddkmvmkiamreikilqlrh 60  
QY 61 PNLVNIIEVFRKRMELHFEYCDHTLLNELEPNPVGADGVIKSVLAQTQLALNECHIH 120  
Db 61 enlvnlleevckkrwlvfefvdhtlldllefpglddqvqvqylqilingigichsh 120  
QY 121 NCIHRDIKPNILITKQIIKICDFEQAQIL-IPCDAYTDYVATRWYRAPELLVGDQYTG 179  
Db 121 nihrdikpenilvsqsvvkkicdfgfarthlaapgevydyvatrwyrapelvlgdvkyg 180  
QY 180 SSVDIWAIGCVFAELLTGTGPLWPKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP 239  
Db 181 kavdvwaigclvtemlmgeplfpgdsldlqlylmrc----- 218  
QY 240 AVASOSAGITGKLIPIHQSIKSNFGFHGISIPEPEDMETLEEKPSDVHPVALNFMKGL 299  
Db 219 -----gnliprhqelfyknvpfagvrpeikeksepterrypkisevvidiakkcl 268  
QY 300 KNPDDRLTCSQLESSEYF--DSF-----OEAQIK--RKARN-----EGNRRRROQ 341  
Db 269 hvdpdkrpfcaellhhdffqmdgfaerfsqelgmkvqkdarnislksksqnrkke 324

RESULT 9  
ID AAE00490 standard; Protein; 187 AA.  
XX AC AAE00490;  
XX DT 19-JUN-2001 (first entry)  
XX DE Human kinase #1.  
XX KW Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder.  
XX OS Homo sapiens.  
XX PN WO200123579-A1.  
XX PD 05-APR-2001.  
XX PF 27-SEP-2000; 2000WO-US26621.  
XX PR 28-SEP-1999; 99US-0156511.  
XX PA (LEXI-) LEXICON GENETICS INC.  
XX PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI; 2001-266166/27.  
XX DR N-PSDB; AAD03812, AAD03818.  
XX CC New isolated human kinase polynucleotide useful for generating  
XX antibodies, as reagents in diagnostic assays and for screening for  
XX compounds useful for treating mental, biological or medical diseases -  
XX PS Disclosure; Page 27-28; 38pp; English.  
XX CC The present sequence is novel human protein (NHP) known  
XX as human kinase. The human kinases share structural  
XX similarity with animal kinases, more particularly serine or  
XX threonine protein kinases. Human kinase cDNA is useful for the  
XX detection of mutant human kinase for the diagnosis of disease,  
XX and also as a therapeutic. It is useful for screening drugs,  
XX effective in the treatment of symptomatic or phenotypic  
XX manifestations perturbing the normal function of NHP in the



body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide construct encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a continuous supply of NHP to the body. Nucleotide constructs encoding functional NHPs are used in gene therapy for the modulation of NHP expression.

Sequence 187 AA;

Query Match 43.5%; Score 791; DB 22; Length 187;  
Best Local Similarity 100.0%; Pred. No. 8.5e-77;  
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGECSGVVFKCRNKTSGVVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60  
|||||  
DB 1 mekyeklaktgecsygvvfkcrnktsgvvavkvvfveseddpvkvkialreirmlkqlkh 60  
|||||

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNGVADGVKSVLWOTLQALNFCIH 120  
|||||  
DB 61 pnlvnlievfrkrkmhlfeycdhtllnelernpngvadgvksvlwqtqlqalnfcih 120  
|||||

QY 121 NCIHRRDIPENILITKGIKICDFGFAQL 151  
|||||  
DB 121 nclhrdipenilitkgiikicdfgfaql 151  
|||||

RESULT 10  
AAE00493  
ID AAE00493 standard; Protein; 198 AA.  
XX  
AC AAE00493;  
XX  
DT 19-JUN-2001 (first entry)  
XX  
DE Human kinase #4.  
XX  
KW Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200123579-A1.  
XX  
PD 05-APR-2001.  
XX  
PF 27-SEP-2000; 2000WO-US26621.  
XX  
PR 28-SEP-1999; 99US-0156511.  
XX  
PA (LEXI-) LEXICON GENETICS INC.  
XX  
PI Doncho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
DR N-PSDB; AAD03815.  
DR WPI; 2001-266166/27.  
DR N-PSDB; AAD03815.  
XX New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases  
XX  
PS Disclosure; Page 31; 38pp; English.  
XX  
XX The present sequence is novel human protein

(NHP) known as human kinase. The human kinases share structural similarity with animal kinases, more particularly serine or threonine protein kinases. Human kinase cDNA is useful for the detection of mutant human kinase for the diagnosis of disease, and also as a therapeutic. It is useful for screening drugs effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying polymorphisms and as primers in amplification assays to detect mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide construct encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a continuous supply of NHP to the body. Nucleotide constructs encoding functional NHPs are used in gene therapy for the modulation of NHP expression.

Sequence 198 AA;

Query Match 43.5%; Score 791; DB 22; Length 198;  
Best Local Similarity 100.0%; Pred. No. 9.3e-77;  
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKYEKLAKTGECSGVVFKCRNKTSGVVAVKVFVESEDDPVVKKIALREIRMLKQLKH 60  
|||||  
DB 1 mekyeklaktgecsygvvfkcrnktsgvvavkvvfveseddpvkvkialreirmlkqlkh 60  
|||||

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLNELERNPNGVADGVKSVLWOTLQALNFCIH 120  
|||||  
DB 61 pnlvnlievfrkrkmhlfeycdhtllnelernpngvadgvksvlwqtqlqalnfcih 120  
|||||

QY 121 NCIHRRDIPENILITKGIKICDFGFAQL 151  
|||||  
DB 121 nclhrdipenilitkgiikicdfgfaql 151  
|||||

RESULT 11  
AAE00126  
ID AAE00126 standard; Protein; 455 AA.  
XX  
AC AAE00126;  
XX  
DT 06-JUN-2000 (first entry)  
XX  
DE Human lost in leukaemia kinase (LLK).  
XX  
KW Lost in leukaemia kinase; LLK; cdc2-related kinase; human; leukaemia;  
KW MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;  
KW chromosome 5q31.1; cytostatic; mitotic index; treatment; prevention;  
KW gene therapy.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Binding-site /label= ATP\_binding\_site  
FT Binding-site /label= Cyclin\_binding\_motif  
FT Region /label= Consensus\_sequence  
FT /note= "Serine/threonine specific kinase sequence"  
FT Domain /label= MAP\_kinase\_activation\_motif  
FT Region /label= Consensus\_sequence  
FT /note= "Serine/threonine specific kinase sequence"

XX PN WO200012719-A1.  
 XX PD 09-MAR-2000.  
 XX PF 31-AUG-1999; 99WO-CA00794.  
 XX PR 31-AUG-1998; 98CA-2243784.  
 XX PR 20-NOV-1998; 98CA-2251249.  
 XX PA (ONTA-) ONTARIO CANCER INST.  
 XX PI Zanke B, Haq R, Randall S, Midmer M;  
 XX DR WPI; 2000-237880/20.  
 XX DR N-PSDB; AAZ51208.  
 XX PT Isolated polynucleotide encoding a lost in leukemia kinase (LLK)  
 XX PT protein, useful for treatment, diagnosis and prevention of leukemia -  
 XX PS Claim 17; Page 51; 69pp; English.  
 XX CC The present amino acid sequence is the human lost in leukaemia kinase  
 CC (LLK), isolated from human lambda gt11 foetal heart genomic library.  
 CC This gene is mapped to chromosome 5q31.1, a region implicated in human  
 CC acute leukaemia. It is expressed strongly in muscle, heart, liver, brain  
 CC and in tissues which have a very low mitotic index. Human LLK has 83%  
 CC amino acid sequence homology to rat LLK beta protein. LLK is closely  
 CC related to cdc2-related kinases, that are putative tumour suppressors  
 CC and to mitogen activated protein kinases (MAPKs). This sequence has  
 CC cytoskeletal activity. It is useful for the treatment, diagnosis and  
 CC prevention of acute leukaemia and is also used in gene therapy.  
 XX CC Sequence 455 AA;

Query Match 41.7%; Score 758.5; DB 21; Length 455;  
 Best Local Similarity 44.0%; Pred. No. 1.1e-72;  
 Matches 150; Conservative 61; Mismatches 91; Indels 39; Gaps 5;  
 QY 1 MEKYEKLAGTGGSGVVFVKCRNKTSGQVAVKKFVESEDDPVVKIALREIRMLKQLKH 60  
 Db 1 memyetlgkvgesgvtvmckhktgqivaikiferyperqgs-vnkiamreikflkqfhh 59  
 QY 61 PNLVNLIEVFRKRKMHVPEYCDHTLLNELERNPNVADGVVKSVMQTLQALNFCCHI 120  
 Db 60 enlvnlievfrqkkihlfvfidhtvldelqhychgleskrkrlyfqliraidylhsh 119  
 QY 121 NCIHRTKPENILITKOGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGTQYG 179  
 Db 120 nihrdikpenilvsggiktdcfgartlaapgdlydvatrwyrapelviktsyg 179  
 QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIITLVETGFRHVDQAGLELLTSSDPP 239  
 Db 180 kpvdilwalgcmilematgpnypipssddilhlhkvk----- 217  
 QY 240 AVASOAGITCKLIPRHQSFKSNFGFFHGISTPEPEDEMTLEKFSVDHPVALNFMKGL 299  
 Db 218 -----gnlsphlqnifskspitagvvlpqvqhpknarkkypkplngliadvhac 267  
 QY 300 KNPDPDRITCSQLESSEYP--DSFOE---AQIKRKARNEGR 335  
 Db 268 qidpadrisssdlhbeyftdrgfiekfmpelkakllqleak 308

RESULT 12  
 AAU03524  
 ID AAU03524 standard; Protein; 591 AA.  
 XX AC AAU03524;  
 XX AC AAU03524;  
 XX DT 12-SEP-2001 (first entry)  
 XX

DE XX Human protein kinase #24.  
 KW XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
 KW XX metabolic disorder; immune related disease; neurological disorder;  
 KW XX neurodegenerative disorder; inflammatory disorder; infectious disease;  
 XX XX reproductive disorder.  
 OS XX Homo sapiens.  
 XX PN WO200138503-A2.  
 XX PD 31-MAY-2001.  
 XX PF 22-NOV-2000; 2000WO-US32085.  
 XX PR 24-NOV-1999; 99US-0167482.  
 XX PA (SUGE-) SUGEN INC.  
 XX PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
 XX PI Flanagan P, Clary B;  
 XX DR WPI; 2001-343950/36.  
 XX DR N-FSDB; AAS06724.  
 XX PT Nucleic acids encoding human kinase polypeptides, useful for preventing  
 XX PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
 XX PT neuronal-associated diseases, and microbial infections -  
 XX PS Claim 7; Figure 2; 433pp; English.  
 XX CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The  
 CC novel protein kinases have been identified as members of the tyrosine  
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides  
 CC encoding protein kinases and the polypeptides may be used in the  
 CC prevention, diagnosis and treatment of diseases associated with  
 CC inappropriate kinase expression. For example, they may be used to treat  
 CC cancers (especially cancers of hematopoietic origin), cardiovascular  
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
 CC immune related diseases (e.g. rheumatoid arthritis), neurological  
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
 CC Additionally, polynucleotides encoding protein kinases may be  
 CC used for gene therapy and as DNA probes in diagnostic assays.  
 CC The protein kinase polypeptides may be used as antigens in the production  
 CC of antibodies against the protein kinases and in assays to identify  
 CC modulators of protein kinase expression and activity.  
 XX CC Sequence 591 AA;

Query Match 41.7%; Score 758.5; DB 22; Length 591;  
 Best Local Similarity 44.0%; Pred. No. 1.6e-72;  
 Matches 150; Conservative 61; Mismatches 91; Indels 39; Gaps 5;  
 QY 1 MEKYEKLAGTGGSGVVFVKCRNKTSGQVAVKKFVESEDDPVVKIALREIRMLKQLKH 60  
 Db 1 memyetlgkvgesgvtvmckhktgqivaikiferyperqgs-vnkiamreikflkqfhh 59  
 QY 61 PNLVNLIEVFRKRKMHVPEYCDHTLLNELERNPNVADGVVKSVMQTLQALNFCCHI 120  
 Db 60 enlvnlievfrqkkihlfvfidhtvldelqhychgleskrkrlyfqliraidylhsh 119  
 QY 121 NCIHRTKPENILITKOGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGTQYG 179  
 Db 120 nihrdikpenilvsggiktdcfgartlaapgdlydvatrwyrapelviktsyg 179  
 QY 180 SSVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIITLVETGFRHVDQAGLELLTSSDPP 239  
 Db 180 kpvdilwalgcmilematgpnypipssddilhlhkvk----- 217  
 QY 240 AVASOAGITCKLIPRHQSFKSNFGFFHGISTPEPEDEMTLEKFSVDHPVALNFMKGL 299



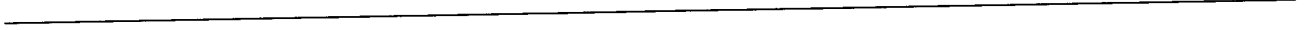
XX Zanke B, Haq R, Randall S, Midmer M;  
 XX WPI: 2000-237880/20.  
 DR N-PSDB; AAZ51206.  
 XX Isolated polynucleotide encoding a lost in leukemia kinase (LLK)  
 PT protein, useful for treatment, diagnosis and prevention of leukemia -  
 XX Claim 15; Page 47; 69pp; English.  
 XX The present amino acid sequence is the alpha isoform of rat lost in  
 CC leukaemia kinase (LLK alpha), isolated from rat brain cDNA library. It  
 CC is expressed strongly in muscle, heart, liver, brain and in tissues which  
 CC have a very low mitotic index. Rat LLK alpha protein has 47% and 44%  
 CC sequence homology to KKTABE and KKTALRE Cdc2-related kinases that are  
 CC putative tumour suppressors and to mitogen activated protein kinases  
 CC (MAPs). This sequence has cytosstatic activity. It is useful for the  
 CC treatment, diagnosis and prevention of acute leukaemia and is also used  
 CC in gene therapy.  
 XX Sequence 505 AA;  
 SQ  
 Query Match 40.5%; Score 737.5; DB 21; Length 505;  
 Best Local Similarity 42.5%; Pred. No. 2.3e-70;  
 Matches 145; Conservative 65; Mismatches 92; Indels 39; Gaps 5;  
 QY 1 MEKYEKLAKTGGSGYGVFKCRNKTSGQVAVKVFVESEDDPVVKKTALREIRMLKQLKH 60  
 Db 1 memyetlgkvgesgygtvmckhktgrivaikfiye-kpeksvniatreakfklqfrh 59  
 QY 61 PNLVNLIEVFRKRKMHILVFECYDHTLNELEPNPNVGVADGVKSVLQWLOALNFCIH 120  
 Db 60 enlvnlievfrgkxkhivfefidhtvldelqhychgeskrirkyfqlraieylnm 119  
 QY 121 NCHRDIKPENILITKQGIKICDFGAQIL-IPGDAYDYVATRWYRAPELVGDTQY 179  
 Db 120 nllhrdikpenilvsqgitklcdfgfartlaapgdvydyvatrwyrapelvlkdtgy 179  
 QY 180 SSVDIWAICGVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP 239  
 Db 180 kpvdiwalgcmilematgpnypipsssdldllhkvikv----- 217  
 QY 240 AVASQAGITGKLTIPRHOSIFKSNFGFHGISIPEPEDMETLEERKFSVDHPVALNFMKGL 299  
 Db 218 -----gnltphlnifskspifagrvlpqvbqphknarkypknglladivhacl 267  
 QY 300 KMPDRLTCSOLLESSYF--DSQF---AQIKRKARNEGR 335  
 Db 268 qdpaerisstdlhhdyftrdgrfiekfipelrakillgeak 308  
 RESULT 15  
 AAB65645  
 ID AAB65645 standard; Protein: 534 AA.  
 XX  
 AC AAB65645;  
 XX  
 DT 27-MAR-2001 (first entry)  
 XX  
 DE Novel protein kinase, SEQ ID NO: 172.  
 XX  
 KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;  
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;  
 KW immune disorder; cardiovascular disease; neurodegenerative disease;  
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.  
 XX Homo sapiens.  
 OS  
 XX WO200073469-A2.  
 PN

XX 07-DEC-2000.  
 XX 26-MAY-2000; 2000WO-US14842.  
 XX 28-MAY-1999; 99US-0136503.  
 XX (SUGE-) SUGEN INC.  
 PA Plowman GD, Martinez R, Whyte D, Sudersanam S;  
 PI WPI: 2001-032161/04.  
 DR N-PSDB; AAF44672.  
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
 CC treating immune-related diseases and disorders, cardiovascular disease,  
 CC neurodegenerative diseases and/or cancers -  
 PT Claim 10; Fig 1; 310pp; English.  
 XX The present sequence is a novel protein kinase. The novel protein kinases  
 CC and the nucleic acids that encode them may be used in the treatment and  
 CC diagnosis of diseases associated with inappropriate kinase expression  
 CC such as immune-related diseases and disorders, cardiovascular disease,  
 CC neurodegenerative diseases and/or cancers. The nucleic acids and  
 CC complementary sequences may also be used as DNA probes in diagnostic  
 CC assays. The kinase polypeptides may be used as antigens in the production  
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
 CC and kinase antagonists may also be used to down regulate kinase  
 CC expression and activity. Diseases related to kinase expression and  
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
 CC disorders, complications of organ transplantation, myocardial infarction,  
 CC immune disorders, cardiomyopathies, strokes, renal failure,  
 CC oxidative stress-related disorders, chronic inflammatory bowel disease,  
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
 CC reproductive disorders.  
 XX Sequence 534 AA;  
 SQ  
 Query Match 36.6%; Score 666; DB 22; Length 534;  
 Best Local Similarity 38.7%; Pred. No. 1.3e-62;  
 Matches 156; Conservative 68; Mismatches 119; Indels 60; Gaps 12;  
 QY 1 MEKYEKLAKTGGSGYGVFKCRNKTSGQVAVKVFVESEDDPVVKKTALREIRMLKQLKH 60  
 Db 1 memyetlgkvgesgygtvmckhktgrivaikfiyerpeqs-vnkiamreikfklqfhh 59  
 QY 61 PNLVNLIEVFRKRKMHILVFECYDHTLNELEPNPNVGVADGVKSVLQWLOALNFCIH 120  
 Db 60 enlvnlievfrgkxkhivfefidhtvldelqhychgeskrirkyfqlraieylnm 119  
 QY 121 N-CHRDIKPENILITKQGIKICDFGAQIL-IPGDAYDYVATRWYRAPELVGDTQY 178  
 Db 120 nvlhrrdikpenilvsqgitklcdfgfartlaapgdlydyvatrwyrapelvlkdtys 179  
 QY 179 GS--SVDIWAICGVFAELLTGQPLWPGKSDVDQLYLI---IRTLVETGFRHVDQAGLELL 233  
 Db 180 gkyvpvdiwalgcmilematgpnypipsssdldllhkvikvxfmpelkakillgeakvnsl 239  
 QY 234 TSSDPPAVASQ-----SAGITGKLIPRHQS-----IFKSNFGFHGISI 271  
 Db 240 ikpesskenelrkderktyvntllsssvlgkeiekpkpkeikvrvikvkggrgdis 299  
 QY 272 PEPDME---TLEKESVDHVA-----LNFMKGC--LKMN-----PD 304  
 Db 300 pkkkeyeglgqdanenvhpmtpdklvtieppnipnstncnglkenphcgsgsvtmpp 359  
 QY 305 DRLTCSQLLESSYFDSFOAQIKRKARNEGRNRQ---QVLP 344  
 Db 360 inltnsnlmaanlssnlftpsvriterakrrtsssgsgqgvp 402

us-09-671-050-10.rag

Fri May 3 11:22:24 2002

Search completed: May 3, 2002, 11:02:15  
Job time: 50 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: May 3, 2002, 11:01:26 ; Search time 26.4 Seconds  
(without alignments)  
1001.234 Million cell updates/sec

Title: US-09-671-050-10  
Perfect score: 1820  
Sequence: 1 MEKYEKLAKTGGSGYGVFK.....RKARNEGRRRQVPLKLS 347

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1157.5	63.6	358	2 S23383	protein kinase (EC
2	1149.5	63.2	376	2 S22745	serine/threonine p
3	565	31.0	305	2 S23382	protein kinase (EC
4	563	30.9	302	1 T50474	protein kinase (EC
5	558.5	30.7	292	2 S40021	protein kinase (EC
6	555	30.5	297	2 A37871	protein kinase (EC
7	554	30.4	298	2 A41227	protein kinase (EC
8	552.5	30.4	294	2 B40444	protein kinase (EC
9	551.5	30.3	294	2 S23095	protein kinase (EC
10	551.5	30.3	294	2 S22440	protein kinase (EC
11	546.5	30.0	294	1 A40444	protein kinase (EC
12	546.5	30.0	294	2 T49271	CELL DIVISION CONT
13	545	29.9	297	2 A29539	protein kinase (EC
14	545	29.9	298	1 A44878	protein kinase (EC
15	542	29.8	297	2 T45977	cyclin-dependent k
16	541	29.7	297	1 S24313	protein kinase (EC
17	540.5	29.7	297	1 S12009	protein kinase cdc
18	540.5	29.7	302	2 B44349	protein kinase (EC
19	540	29.7	302	1 A44349	protein kinase (EC
20	539	29.6	303	1 S06011	protein kinase (EC
21	537	29.5	288	1 S42566	protein kinase (EC
22	536.5	29.5	311	2 S36619	protein kinase (EC
23	535.5	29.4	294	1 J02243	protein kinase (EC
24	535.5	29.4	294	1 S57928	protein kinase (EC
25	535.5	29.4	294	1 S42049	protein kinase (EC
26	532	29.2	301	1 S19209	protein kinase (EC
27	531	29.2	292	1 A46365	protein kinase (EC
28	530.5	29.1	294	1 S31332	protein kinase (EC
29	529	29.1	292	1 I49592	protein kinase (EC

30	529	29.1	292	1 A45091	protein kinase (EC
31	527.5	29.0	308	1 S53538	protein kinase (EC
32	527	29.0	307	1 A36074	protein kinase (EC
33	527	29.0	301	1 S42101	protein kinase (EC
34	526.5	28.9	291	2 S23386	protein kinase (EC
35	526	28.9	293	2 JE0374	cyclin-dependent k
36	526	28.9	294	2 S51008	protein kinase (EC
37	526	28.9	296	2 S24386	protein kinase (EC
38	524	28.8	302	1 OKB85	protein kinase PHO
39	521.5	28.7	302	2 T17115	CDK-activating pro
40	519.5	28.5	346	2 A54820	protein kinase (EC
41	518	28.5	292	2 S22441	M015/CDK-activatin
42	517.5	28.4	346	2 A56231	protein kinase (EC
43	516.5	28.4	314	1 S12007	protein kinase (EC
44	514	28.2	293	2 T02922	protein kinase (EC
45	514	28.2	346	1 I78840	protein kinase (EC

ALIGNMENTS

RESULT 1  
S23383  
protein kinase (EC 2.7.1.37) cdc2-related KKIALRE - human  
C:Species: Homo sapiens (man)  
C>Date: 15-Oct-1994 #sequence\_revision 15-Oct-1994 #text\_change 18-Jun-1999  
C:Accession: S23383; S22744  
R:Meyerson, N.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.;  
EMBO J. 11, 2909-2917, 1992  
A:Title: A family of human cdc2-related protein kinases.  
A:Reference number: S23382; MUID:92347325  
A:Accession: S23383  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-358 <MEY>  
A:Cross-references: EMBL:X66358; NID:g36614; PIDN:CAA47002.1; PID:g36615  
C:Superfamily: Kinase-related transforming protein; protein kinase homology  
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F:3-278/Domain: protein kinase homology <KIN>  
F:11-19/Region: protein kinase ATP-binding motif  
F:34,52,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match	63.6%	Score 1157.5;	DB 2;	Length 358;
Best Local Similarity	64.7%	Pred. No. 2.3e-46;		
Matches 211;	Conservative 41;	Mismatches 41;	Indels 33;	Gaps 2;
QY 1	MEKYEKLAKTGGSGYGVFKCRNKTSGQVAVKFKFVESEDDPVVKKIALREIRMLKQLKH	60		
Db 2	MEKYEKIGKIGGSGYGVFKCRNKTSGQVAVKFKFVESEDDPVVKKIALREIRMLKQLKH	61		
QY 61	PNLVNLTIEVFRKKRMHLVFEYCDHTLLNELERNPNGVADGVTKSVLWOTLQALNFCIH	120		
Db 62	PNLVNLTIEVFRKKRMHLVFEYCDHTLVHLDYQYGVPEHLVKSITWOTLQAVNFCCHK	121		
QY 121	NCIHRDKPENILITKOGIICDFGFAQILI-PGDVATDYVATRWYRAPELLVGDTOYG	179		
Db 122	NCIHRDKPENILITKOGIICDFGFAQILI-PGDVATDYVATRWYRAPELLVGDTOYG	181		
QY 180	SSVDIWAIGCVFAELITGQPLWPKSDVDOLYLITLTVETGFRHVDQAGLELLTSSDPP	239		
Db 182	PPVDVWAIGCVFAELITGQPLWPKSDVDOLYLITLTVETGFRHVDQAGLELLTSSDPP	219		
QY 240	AVASQAGITGKILPHQSIKSGFFHGISIPEPEDMETLFEKFSVDHVPALNFMKGL	299		
Db 220	-----GDLIPHQVQFSTNQYFSGVKIPDPEDMEPLKLPNITSYALGGLKGL	269		
QY 300	KMNPDDRLTCSQLESSEYFDSFQEAQ	325		
Db 270	HMDPTERTUCQLLHPYFENIRELE	295		
RESULT 2				

S22745

serine/threonine protein kinase KKIALRE (EC 2.7.1.1) - human  
 C:Species: Homo sapiens (man)  
 C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 19-Dec-1997  
 C:Accession: S22745  
 R:Meyserson, M.L.  
 submitted to the EMBL Data Library, May 1992  
 A:Reference number: S22743  
 A:Accession: S22745  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-376 <MEY>  
 A:Cross-references: EMBL:X66359  
 C:Genetics:  
 A:introns: 152/3; 170/3  
 C:Superfamily: Kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
 F:3-296/Domain: protein kinase homology <KIN>  
 F:11-19/Region: protein kinase ATP-binding motif

Query Match 63.2%; Score 1149.5; DB 2; Length 376;  
 Best Local Similarity 61.3%; Pred. No. 5.4e-46;  
 Matches 211; Conservative 41; Mismatches 41; Indels 51; Gaps 2;  
 QY 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGOVAVKKFVESEDDDPVVKKIALREIRMLKQKH 60  
 DB 2 MEKYEKIGKIGEGYGVVFKCRNKTSGOVAVKKFVESEDDDPVVKKIALREIRMLKQKH 61  
 QY 61 PNLVNLIEVFRKKMHLVFEYCDHTLLNELRNPNMGVADGVKISVLWQTLQALNFCVTH 120  
 DB 62 PNLVNLIEVFRKKMHLVFEYCDHTLLNELRNPNMGVADGVKISVLWQTLQALNFCVTH 121  
 QY 121 NCIRHDKIPENILITKOGIIKICDFGFAQLI-----PGDAYTDV 161  
 DB 122 NCIRHDKIPENILITKHSVIKLCDFGFAQLIFXPQSAACVFCPSITGXTGPDSDYTDV 181  
 QY 162 ATRWRAPELLVGDYQSSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRITLVETG 221  
 DB 182 ATRWTRPELLVGDYQGPVVDVWAIGCVFAELLTGQPLWPKSDVDQLYLIIRITLVETG 221  
 QY 222 FRHVDQAGLELLTSSDPPAVASQASAGITGKLIPIHQSIKSNFGFHGISIPPEP-DMETLEEKFSVDVHPVAL 292  
 DB 238 -----GDLIPRQQVFTSNQYFSGVKIPDPEDMEPLE 269  
 QY 282 EKFSVDVHPVALNFMKGLKMPDRLTCSQLLESYFDSFQEAQ 325  
 DB 270 LKFPNISYPALGLLKGCLHMDPTERLTCQLLHPYFENIREIE 313

RESULT 3

S23382  
 protein kinase (EC 2.7.1.37) cdk3 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 18-Jun-1993 #sequence\_revision 18-Jun-1993 #text\_change 18-Jun-1999  
 C:Accession: S23382; S22743  
 R:Meyserson, M.; Ender, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; Tsai  
 EMBO J. 11, 2509-2517, 1992  
 A:Title: A family of human cdc2-related protein kinases.  
 A:Reference number: S23382; MUID:92347325  
 A:Accession: S23382  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-305 <MEY>  
 A:Cross-references: EMBL:X66357; NID:g36612; PIDN:CAA47001.1; PID:g36613  
 C:Genetics:  
 A:Gene: GDB:CDK3  
 A:Cross-references: GDB:283456  
 A:Map position: 12q13-12q13  
 C:Superfamily: Kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase  
 F:2-255/Domain: protein kinase homology <KIN>  
 F:10-18/Region: protein kinase ATP-binding motif

F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 31.0%; Score 565; DB 2; Length 305;  
 Best Local Similarity 36.8%; Pred. No. 1.6e-19;  
 Matches 124; Conservative 66; Mismatches 99; Indels 48; Gaps 9;

QY 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGOVAVKKF---VESEDDDPVVKKIALREIRMLKQ 57  
 DB 1 MDMFKVKEIGEGYGVVYKAKNRETQGLVALKIRLDLEMGVP---STAIREISLKE 57  
 QY 58 LKHPNLVNLIEVFRKKMHLVFEYCDHTLLNELRNPNMGVADGVKISVLWQTLQALNFC 116  
 DB 58 LKHPNLVNLVLDVHNERKLYLVFEFLSODLAKKYMDSTPGSELPLMLIKSYLQQLGVSF 117  
 QY 117 CHHCICIRHDKIPENILITKOGIIKICDFGFAQLI-IPGDAYTDVATWRAPELLVGD 175  
 DB 118 CHSHRVIRHDKIPONLLINELGAIKLADFGARAFGVPLRTVTHEVVTWRAPELLG 177  
 QY 176 TOYGSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRITLVETGFRHVDQAGLELLTS 235  
 DB 178 KFYTTAVDINSIGCIFAEWVTRKALFPGDSEIDQLERFRL----- 219  
 QY 236 SDPPAVASOSA--GITGKLIPIHQSIKSNFGFHGISIPPEP-DMETLEEKFSVDVHPVAL 292  
 DB 220 ---GTPSEDTPGVTI--QLPDYKGSF-----PKWTRKGLLEEIVNLEPEGR 260  
 QY 293 NFMKGLKMPDRLTCSQLLESYFDSFQEAQIKRK 329  
 DB 261 DLLMQLLOYDPSORITAKTALAHFYFSSPEPPSPAAQ 297

RESULT 4

I50474  
 protein kinase (EC 2.7.1.37) cdc2 [similarity] - goldfish  
 C:Species: Carassius auratus (goldfish)  
 C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
 C:Accession: I50474  
 R:Kajiuura, H.; Yamashita, M.; Katsu, Y.; Nagahama, Y.  
 Dev. Growth Differ. 35, 647-654, 1993  
 A:Title: Isolation and characterization of goldfish cdc2, a catalytic component of ma  
 A:Reference number: I50474  
 A:Accession: I50474  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-302 <KAJ>  
 A:Cross-references: GB:D17758; NID:g471097; PIDN:BA04605.1; PID:g471098  
 C:Superfamily: Kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphotransferase  
 F:2-256/Domain: protein kinase homology <KIN>  
 F:10-18/Region: protein kinase ATP-binding motif  
 F:33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.9%; Score 563; DB 1; Length 302;  
 Best Local Similarity 41.2%; Pred. No. 1.9e-19;  
 Matches 117; Conservative 60; Mismatches 71; Indels 36; Gaps 5;

QY 1 MEKYEKLAKTGGEGYGVVFKCRNKTSGOVAVKKF-VESEDDDPVVKKIALREIRMLKQ 59  
 DB 1 MDDYLKIEKIGEGYGVVYKGRNKTGQVAVKMKIRLESEERG-VPSTAVREISLKELO 59  
 QY 60 HPNLVNLIEVFRKKMHLVFEYCDHTLLNELRNPNMGVADGVKISVLWQTLQALNFC 117  
 DB 60 HPNVRLLDVLMQESKLYLVFEFLSMDLKYLDISIPSGFMDPMLKSYLQILLEGILFC 119  
 QY 118 HHHCICIRHDKIPENILITKOGIIKICDFGFAQLI-IPGDAYTDVATWRAPELLVGD 176  
 DB 120 HCRVRLHRLKPNQLLIDNKGVIKLADFGARAFGVPLRVVTHEVVTWRAPELLG 179  
 QY 177 QYGSVDIWAIGCVFAELLTGQPLWPKSDVDQLYLIIRITL----- 217  
 DB 180 RYSTPVDVMSIGTIFAELATKPLFHGUSEIDQLFRIFRTIGTPNPNVWPDVPSLPDYKN 239





A:Reference number: A41227; MUID:92020980

A:Accession: A41227

A:Molecule type: mRNA

A:Residues: 1-298 <IN>

A:Cross-references: GB:M68520; NID:g180177; PIDN:AAA35667.1; PID:g180178

R:Tsal, L.H.; Harlow, E.; Meyerson, M.

Nature 353, 174-177, 1991

A:Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1A-

A:Reference number: S17873; MUID:91367262

A:Accession: S17873

A:Molecule type: mRNA

A:Residues: 1-298 <TSA>

A:Cross-references: GB:X62071; NID:g312802; PIDN:CAA43985.1; PID:g312803

R:Elledge, S.J.; Spottswood, M.R.

EMBO J. 10, 2653-2659, 1991

A:Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc28

A:Reference number: S16520; MUID:91330891

A:Accession: S16520

A:Molecule type: preliminary

A:Residues: 1-176, 'S', 178-298 <ELL>

A:Cross-references: EMBL:X61622; NID:g29848; PIDN:CAA43807.1; PID:g29849

C:Genetics: GDB:CDK2

A:Gene: GDB:CDK2

A:Cross-references: GDB:128984; OMIM:116953

A:Map position: 12q13-12q13

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine

F:2-255/Domain: protein kinase homology <KIN>

F:10-18/Region: protein kinase homology <KIN>

F:14,160/Binding site: phosphate (Thr) (covalent) #status predicted

F:15/Binding site: phosphate (Tyr) (covalent) #status predicted

F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.4%; Score 554; DB 2; Length 298;

Best Local Similarity 38.4%; Pred. No. 4.9e-19;

Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGOVAVKFFVESEDDPVVKIALREIRMLKOLKH 60

Db 1 MEQYKVKIKGEGTYGVYKALDKDTANETIALKIRLEQEDRGVPSTAIRSILAKEMNH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVKISVLWQTLQAL 119

Db 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVKISVLWQTLQAL 119

QY 120 HNCIHRDIKIPENILITKOGIILKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDQY 178

Db 121 HRLVRLDKPQNLITTEGAIKLAFLARAFGVPVRYTHEVVTWLYRAPELLVGDQY 180

QY 179 GSSVDIWAICGVFAELLTGQPLWPKSDVDQLYLIIRTL----- 217

Db 181 STAVDWSLGCIFADWVTRALFGDSEIDQLFRIETLGTDPDVVWPGVSMDPYKPSF 240

QY 218 ---VETGFRHV---DQAGLELLTSS---DPPAVASQASAGITCKLIPRHOSIFKSNFFH 267

Db 241 PKWARQDFSKVVPPLDEGRSLLSQMLHDPNKRISAKAAL-----AHPFEQ 287

QY 268 GISPEP 274

Db 288 DVTKPVP 294

RESULT 8

B40444

protein kinase (EC 2.7.1.37) cdc2 homolog B - maize (fragment)

C:Species: Zea mays (maize)

C:Date: 14-Feb-1992 #sequence\_revision 14-Feb-1992 #text\_change 28-Feb-1997

C:Accession: B40444

R:Colasanti, J.; Tyers, M.; Sundaresan, V.

Proc. Natl. Acad. Sci. U.S.A. 88, 3377-3381, 1991

A:Title: Isolation and characterization of cDNA clones encoding a functional p34 (cdc2)

A:Reference number: A40444; MUID:91195354

A:Accession: B40444

A:Molecule type: preliminary

A:Residues: 1-294 <COL>

A:Cross-references: GB:M60526

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphotransferase

F:2-256/Domain: protein kinase homology <KIN>

F:10-18/Region: protein kinase ATP-binding motif

F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.4%; Score 552.5; DB 2; Length 294;

Best Local Similarity 35.4%; Pred. No. 5.7e-19;

Matches 119; Conservative 66; Mismatches 98; Indels 53; Gaps 9;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGOVAVKFFVESEDDPVVKIALREIRMLKOLKH 60

Db 1 MEQYKVKIKGEGTYGVYKALDKDTANETIALKIRLEQEDRGVPSTAIRSILAKEMNH 60

QY 61 PNLVNLIEVFRKRKMHVFEYCDHTLLN-----ELERNPNGVADGVKISVLWQTLQAL 114

Db 61 GNVRLHDVHSEKRIYLVFEYLDLKLKPMWSDSCPEFAKNPT-----LKSYYLQILRGV 115

QY 115 NFCHTHNCIHRDIKIPENILITKQ-GIIKICDFGAQIL-IPGDAYTDYVATRWYRAPELL 172

Db 116 AYCHSHRVLRDLPQNLITTEGAIKLAFLARAFGVPVRYTHEVVTWLYRAPELL 175

QY 173 VEDTQYSSVDIWAICGVFAELLTGQPLWPKSDVDQLYLIIRTL---VETGFRHVDQAG 229

Db 176 LGARQYTPVDVWVSGCIFAEWVQKPLFGDSEIDELFKIFRVLGTDPNEQGW----- 228

QY 230 LELLTSDPPAVASQASAGITCKLIPRHOSIFKSNFFHGHSIPPEMETLEKFSVDVHP 289

Db 229 -----PGVSCLPDFKTA--FPRWQA-----ODLATI---VPNLEP 258

QY 290 VALNFMKGLKMLRYEPVSKRITARQALEHEFYKDLQV 325

Db 259 AGDLLSKMLRYEPVSKRITARQALEHEFYKDLQV 294

RESULT 9

S23095

protein kinase (EC 2.7.1.37) cdc2 - Arabidopsis thaliana

N:Alternate names: cdc2 protein homolog; CDC2a protein; cell division control protein

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 22-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 20-Jun-2000

C:Accession: S23095; A48984; JQ1337; JQ0967; S18202

R:Imajuku, Y.; Hirayama, T.; Endoh, H.; Oka, A.

FEBS Lett. 304, 73-77, 1992

A:Title: Exon-intron organization of the Arabidopsis thaliana protein kinase genes CD

A:Reference number: S23095; MUID:92316202

A:Accession: S23095

A:Molecule type: DNA

A:Residues: 1-294 <INA>

A:Cross-references: EMBL:D10850; NID:g217848; PIDN:BAA01623.1; PID:g217849

R:Inze, D.; Ferreira, P.; Hemerly, A.; Van Montagu, M.

Biochem. Soc. Trans. 20, 80-84, 1992

A:Title: Control of cell division in plants.

A:Reference number: A48984; MUID:92339744

A:Accession: A48984

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-294 <INZ>

A:Experimental source: flower

A:Note: Sequence extracted from NCBI backbone (NCBIP:109461)

R:Hirayama, T.; Imajuku, Y.; Anai, T.; Matsui, M.; Oka, A.

Gene 105, 159-165, 1991

A:Title: Identification of two cell-cycle-controlling cdc2 gene homologs in Arabidops

A:Reference number: JQ1337; MUID:92039027

A:Accession: JQ1337

us-09-671-050-10.rpr

Fri May 3 11:22:25 2002

F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.3%; Score 551.5; DB 2; Length 294;  
 Best Local Similarity 35.5%; Pred. No. 6.3e-19;  
 Matches 115; Conservative 71; Mismatches 95; Indels 43; Gaps 8;

A:Molecule type: mRNA  
 A:Residues: 1-294 <HF>  
 A:Cross-references: EMBL:X57839; NID:g16218; PIDN:CAAA0971.1; PID:g16219  
 R:Perreira, P.C.G.; Hemery, A.S.; Villarroel, R.; Van Montagu, M.; Inze, D.  
 Plant Cell 3, 531-540, 1991  
 A:Title: The Arabidopsis functional homolog of the p34cdc2 protein kinase.  
 A:Reference number: JQ0967; MUID:93005715  
 A:Accession: JQ0967  
 A:Molecule type: mRNA  
 A:Residues: 1-294 <FE>  
 A:Cross-references: GB:S45387; NID:g257373; PIDN:AB23643.1; PID:g257374  
 A:Experimental source: flower  
 A:Comment: The protein is a key component of the eukaryotic cell cycle.  
 C:Genetics:  
 A:Gene: cdc2  
 A:Introns: 3/3; 65/2; 105/3; 163/3; 218/2; 247/3; 265/3  
 C:Function:  
 A:Description: phosphotransferase; protein kinase; required for G1 to S-phase transition  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase  
 F:2-256/Domain: protein kinase homology <KIN>  
 F:10-18/Region: protein kinase ATP-binding motif  
 F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.3%; Score 551.5; DB 2; Length 294;  
 Best Local Similarity 35.5%; Pred. No. 6.3e-19;  
 Matches 115; Conservative 71; Mismatches 95; Indels 43; Gaps 8;

QY 1 MEKYEKAKTGGSGVGVFKRNKTSQGVAVKFESEDDPVVKKIALREIRMLKOLKH 60  
 DB 1 MDQYKVEKIGEGTYGVVYKARDKVTNETIALKKIRLEQDEGVSTAIRISLLKEMOH 60

QY 61 PNLVNLIEVFRKRMHLLVFEYCDHTLLN-----ELERNPNGVADGVKSVLWQTLQAL 119  
 DB 61 SNVLQDVVHSEKRLVFEYLDLKKHMDSTDFSKLHMIXLYQLIRGIAYCHS 120

QY 120 HNCIHRDKIPENILITKO-GIIKICDFGAQIL-IPGDATDVVATRWYRAPELLVGDTQ 177  
 DB 121 HRVLHRLDKPQNLIDRRTNALADFGARAFGIPVTFTEHVVTLMYRAPELLLSGH 180

QY 178 YGSSVDIWAIGCVFAELLTGQPLWPKGSDVDQLYLIIR---TLVETGFHVDQAGLELLT 234  
 DB 181 YSTPVDIWSVGICFAEMISQKPLFGDSEIDQLFKIRIMGTPTVETWR----- 229

QY 235 SSDPPAVASQSGAGITGKLIIPHQSIKSNFGFHGISIPEPEMETLEEFSDVHPVALNF 294  
 DB 230 -----GVTS--LPDYKSAPFK-----WKPTDLETf---VPLNDPDGVDL 263

QY 295 MKGCKLMNPDDRLTCSOLLESSYF 318  
 DB 264 LSKMLMDPTKRINARAALAEHYF 287

RESULT 10  
 S22440  
 protein kinase (EC 2.7.1.37) cdc2 homolog 1 - rice  
 C:Species: Oryza sativa (rice)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 18-Jun-1999  
 C:Accession: S22440  
 R:Hashimoto, J.; Hirabayashi, T.; Hayano, Y.; Hata, S.; Ohashi, Y.; Suzuka, I.; Utsugi, M.; Gen. Genet. 233, 10-16, 1992  
 A:Title: Isolation and characterization of cDNA clones encoding cdc2 homologues from Ory  
 A:Reference number: S22440; MUID:92293101  
 A:Accession: S22440  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-294 <HAS>  
 A:Cross-references: EMBL:X60374; NID:g20342; PIDN:CAA42922.1; PID:g20343  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase  
 F:2-256/Domain: protein kinase homology <KIN>  
 F:10-18/Region: protein kinase ATP-binding motif

F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.3%; Score 551.5; DB 2; Length 294;  
 Best Local Similarity 34.8%; Pred. No. 6.3e-19;  
 Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;

QY 1 MEKYEKAKTGGSGVGVFKRNKTSQGVAVKFESEDDPVVKKIALREIRMLKOLKH 60  
 DB 1 MEQYKVEKIGEGTYGVVYKARDKVTNETIALKKIRLEQDEGVSTAIRISLLKEMHH 60

QY 61 PNLVNLIEVFRKRMHLLVFEYCDHTLLN-----ELERNPNGVADGVKSVLWQTLQAL 114  
 DB 61 GNIVRLHDVTHSEKRLVFEYLDLKKFMSQCFEFAKNPT-----LKSXYLQILRGV 115

QY 115 NFCIHNCIHRDKIPENILITKO-GIIKICDFGAQIL-IPGDATDVVATRWYRAPELL 172  
 DB 116 AYCHSHRVLHRLDKPQNLIDRRTNALADFGARAFGIPVTFTEHVVTLMYRAPELL 175

QY 173 YGSSVDIWAIGCVFAELLTGQPLWPKGSDVDOLYLIIRLIVETGFRHVDQAGLEL 232  
 DB 176 LGSQYQSTPVDIWSVGICFAEMVQKPLFGDSEIDQLFKIRVLGTP----- 223

QY 233 LTSDPPAVASQSGAGITGKLIIPRHOSIFKSNFGFHGISIPEPEMETLEEFSDVHPVAL 292  
 DB 224 -NEQSWPGVSS-----LPDYKSAPFK-----WQADLIATf---VPTLDPAQL 261

QY 293 NFMKCKLMNPDDRLTCSOLLESSYFDSFOEAQ 325  
 DB 262 DLLSKMLRYEPNKRITARQALEHEYFKDLENVQ 294

RESULT 11  
 A40444  
 protein kinase (EC 2.7.1.37) cdc2 homolog A - maize  
 C:Species: Zea mays (maize)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A40444  
 R:Colasanti, J.; Myers, M.; Sundaresan, V.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 3377-3381, 1991  
 A:Title: Isolation and characterization of cDNA clones encoding a functional p34 (cdc2  
 A:Reference number: A40444; MUID:91195354  
 A:Accession: A40444  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-294 <COL>  
 A:Cross-references: GB:M60526  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphotransferase  
 F:2-256/Domain: protein kinase homology <KIN>  
 F:10-18/Region: protein kinase ATP-binding motif  
 F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 30.0%; Score 546.5; DB 1; Length 294;  
 Best Local Similarity 34.1%; Pred. No. 1.1e-18;  
 Matches 117; Conservative 66; Mismatches 93; Indels 67; Gaps 9;

QY 1 MEKYEKAKTGGSGVGVFKRNKTSQGVAVKFESEDDPVVKKIALREIRMLKOLKH 60  
 DB 1 MEQYKVEKIGEGTYGVVYKARDKVTNETIALKKIRLEQDEGVSTAIRISLLKEMNH 60

QY 61 PNLVNLIEVFRKRMHLLVFEYCDHTLLN-----ELERNPNGVADGVKSVLWQTLQAL 114  
 DB 61 GNIVRLHDVHSEKRLVFEYLDLKKFMSQCFEFAKNPT-----LKSXYLQILHGV 115

QY 115 NFCIHNCIHRDKIPENILITKO-GIIKICDFGAQIL-IPGDATDVVATRWYRAPELL 172  
 DB 116 AYCHSHRVLHRLDKPQNLIDRRTNALADFGARAFGIPVTFTEHVVTLMYRAPELL 175

QY 173 YGSSVDIWAIGCVFAELLTGQPLWPKGSDVDOLYLIIRLIVETGFRHVDQAGLEL 232  
 DB 176 LGSQYQSTPVDIWSVGICFAEMVQKPLFGDSEIDQLFKIRVLGTP----- 221

QY 233 LTSSDPPAVASQAGITGKLIPRHOSIFKNGFFHGIS-IP-----EPEDMETLEE 282  
 Db 222 -----TPNEQS-----WPGVSCLPDFKTAFFRWAQDLATV-- 252  
 QY 283 RFSOVHPVALNFMKGLKMNPPDRLTCSQLLESSYFDSFQFAQ 325  
 Db 253 -VPNLDPAGLDLSKMLRYEFSKRITARQALEHEYFKDLVVQ 294

RESULT 12

CELL DIVISION CONTROL PROTEIN 2 HOMOLOG A - Arabidopsis thaliana  
 N:Alternate names: protein T21J18.20  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jun-2000  
 C:Accession: T49271  
 R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Rudd, S.  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: 225021  
 A:Accession: T49271  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-294 <RTE>  
 A:Cross-references: EMBL:AL132963; GSPDB:GN00061; ATSP:T21J18.20  
 A:Experimental source: cultivar Columbia; BAC clone T21J18  
 C:Genetics:  
 A:Gene: ATSP:T21J18.20  
 A:Map position: 3  
 A:Introns: 3/3; 67/3; 105/3; 163/3; 218/2; 247/3; 265/3  
 C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 30.0%; Score 546.5; DB 2; Length 294;  
 Best Local Similarity 35.2%; Pred. No. 1.1e-18;  
 Matches 114; Conservative 71; Mismatches 96; Indels 43; Gaps 8;

QY 1 MEKYEKLAKTGEISYGVFKCRNKTSGOVAVKKFVESEDDPVVKKLTALREIMLKQLK 60  
 Db 1 MDQYEKIEKIGETGYGVYKARDVNETALAKITLQDEGVPSTAIRLSLKEMQH 60  
 QY 61 PNLVNLIEVFRKRKMLHVFECYDHTLLNELERNPG--VADGVKSVLMQTLQALNFC 119  
 Db 61 SNIKYDDVHSEKRLYLVEYLDLKLKHMDSPTDFSKDLMIKTYLQILRGIAYCHS 120  
 QY 120 HNCIHRDKPENILITKO-GIHKICDFGFAQL-IPGDAYTDYVATRWVRAPELLVGD 177  
 Db 121 HRVLRDLKPNLLIDRRNTSLKLDADGLARAFGIPVTFTHVWTLWTRAPEILLGSHH 180  
 QY 178 YGSSVDIWAIGCVFAELITGOLPWPKSDVDQLYLIIR---TLVETGFRHVDQAGLELLT 234  
 Db 181 YSTPVDIWSVGCIFARMISQKPLFGDSEIDQLKIFRMGTPYEDTWR----- 229  
 QY 235 SSDPPAVASQAGITGKLIPRHOSIFKNGFFHGISTPEDIPEEDMETLEEFSDVHPVALNF 294  
 Db 230 -----GVTS--LPDYKSAPPK-----WKPTDLETF---VPLNDPDGVDL 263  
 QY 295 MKGKLMNPPDRLTCSQLLESSYF 318  
 Db 264 LSKWLLMDPTKRINARAALHEIF 287

RESULT 13

protein kinase (EC 2.7.1.37) cdc2 - human  
 N:Alternate names: cell division control protein 2 (CDC2)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 21-Jul-2000  
 C:Accession: A29539  
 R:Lee, M.G.; Nurse, P.  
 Nature 327, 31-35, 1987  
 A:Title: Complementation used to clone a human homologue of the fission yeast cell cycle  
 A:Reference number: A29539; MUID:87201915

A:Accession: A29539  
 A:Molecule type: mRNA  
 A:Residues: 1-297 <LEE>  
 A:Cross-references: GB:X05360; NID:g29838; PIDN:CAA28963.1; PID:g29839  
 C:Genetics:  
 A:Gene: GDB:CDC2  
 A:Cross-references: GDB:119052; OMIM:116940  
 A:Map position: 10q21.1-10q21.1  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; ser  
 F:2-256/Domain: protein kinase homology <KIN>  
 F:10-18/Region: protein kinase ATP-binding motif  
 F:14-16/Binding site: phosphate (Thr) (covalent) #status predicted  
 F:15/Binding site: phosphate (Tyr) (covalent) #status predicted  
 F:33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 29.9%; Score 545; DB 2; Length 297;  
 Best Local Similarity 40.2%; Pred. No. 1.3e-18;  
 Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;

QY 1 MEKYEKLAKTGEISYGVFKCRNKTSGOVAVKKFVESEDDPVVKKLTALREIMLKQLK 59  
 Db 1 MEDYTKIEKIGETGYGVYKGRKTTGOVAMKKIRLESEEG-VPSTAIRLSLKELR 59  
 QY 60 HENLVNLIEVFRKRKMLHVFECYDHTLLNELERNPG--VADGVKSVLMQTLQALNFC 117  
 Db 60 HENIVSLQDVLMDQSRLLYLFELFSLMDLKKYLDSTPPGQYMDSSLVAKSYLYQLQGI VFC 119  
 QY 118 HHCNHRDKPENILITKGIKICDFGFAQL-IPGDAYTDYVATRWVRAPELLVGD 176  
 Db 120 HSRVLRDLKPNLLIDRRNTSLKLDADGLARAFGIPVTFTHVWTLWTRAPEILLG 179  
 QY 177 YGSSVDIWAIGCVFAELITGOLPWPKSDVDQLYLIIRTL----- 217  
 Db 180 RYTPVDIWSIGITFAELATKPLFGDSEIDQLFRIFRALGTNNVWPEVESLDQYKN 239  
 QY 218 -----VETGFRHVDQAGLELLTSS---DPPAVASQAGITGKLIPRH 256  
 Db 240 TFPKWKPGSLASHVNLNDENGLDLSKMLIYDP-----AKRISGKMLNH 284

RESULT 14

protein kinase (EC 2.7.1.37) cdk2 [validated] - goldfish  
 C:Species: Carassius auratus (goldfish)  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
 C:Accession: A44878  
 R:Hirai, T.; Yamashita, M.; Yoshikuni, M.; Tokumoto, T.; Kajiuura, H.; Sakai, N.; Naga  
 Dev. Biol. 152, 113-120, 1992  
 A:Title: Isolation and characterization of goldfish cdk2, a cognate variant of the ce  
 A:Reference number: A44878; MUID:92331802  
 A:Accession: A44878  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-298 <HIR>  
 A:Cross-references: GB:S40289; NID:g251619; PIDN:AAB22550.1; PID:g251620  
 A:Experimental source: oocyte  
 A:Note: sequence extracted from NCBI backbone (NCBIN:108782, NCBI:108783)  
 C:Genetics:  
 A:Gene: cdk2  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; ser  
 F:2-255/Domain: protein kinase homology <KIN>  
 F:10-18/Region: protein kinase ATP-binding motif  
 F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match

Best Local Similarity 29.9%; Score 545; DB 1; Length 298;  
 Matches 102; Conservative 46; Mismatches 67; Indels 2; Gaps 2;  
 QY 1 MEKYEKLAKTGEISYGVFKCRNKTSGOVAVKKFVESEDDPVVKKLTALREIMLKQLK 60

RESULT . 15

Query Match

Query Match	Best Local Similarity	Pred. No. 1.7e-18;	Indels	Gaps
117. Conservative	40.2%	57; Mismatches	75;	42; Gaps

[illegible]

Search completed: May 3, 2002, 11:03:16  
Job time: 110 sec



P48609 drosophila  
Q38772 antirrhinum  
P51166 xenopus lae  
P50613 homo sapien  
P29619 oryza sativ  
P17157 saccharomyc  
Q03147 mus musculu  
P23573 drosophila  
P97377 mus musculu  
Q92241 kluyveromyc  
P34556 caenorhabdi  
P54665 trypanosoma

34 525 28.8 294 1 CDK5\_DROME  
35 521.5 28.7 294 1 CC2A\_ANTMA  
36 521 28.6 292 1 CDK5\_XENLA  
37 519.5 28.5 346 1 CDK7\_HUMAN  
38 518 28.5 292 1 CC22\_ORYSA  
39 518 28.5 305 1 PHB5\_YEAST  
40 517.5 28.4 346 1 CDK7\_MOUSE  
41 516.5 28.4 314 1 CC2C\_DROME  
42 516 28.4 346 1 CDK2\_MOUSE  
43 512 28.1 304 1 PHB5\_KLUILA  
44 510 28.0 332 1 CC2\_CAEEL  
45 510 28.0 345 1 CC22\_TRYBB

ALIGNMENTS

RESULT 1  
KKIA\_HUMAN  
ID KKIA\_HUMAN STANDARD; PRT; 358 AA.  
AC Q00532;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE SERINE/THREONINE-PROTEIN KINASE KIALRE (EC 2.7.1.-) (CYCLIN-DEPENDENT  
DE KINASE-LIKE 1).  
GN CDKL1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92347325; PubMed=1639063;  
RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,  
RT Harlow E., Tsai L.-H.  
RL EMBO J. 11:2909-2917(1992).  
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC CDC2/CDKX SUBFAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration -  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
EMBL; X66358; CAA47002.1; -  
DR EMBL; X66359; CAA47002.1; JOINED.  
DR PIR; S22744; S22744.  
DR PIR; S22745; S22745.  
DR PIR; S23383; S23383.  
DR HSSP; P24941; LAQ1.  
DR MIM; 603441; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_kin\_actsite.  
DR Pfam; PF00069; pkinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR Transferrase; Serine/threonine-protein kinase; ATP-binding.  
KW DOMAIN 5 288 PROTEIN KINASE.  
FT NP\_BIND 11 19 ATP (BY SIMILARITY).  
FT BINDING 34 34 ATP (BY SIMILARITY).  
FT ACT\_SITE 127 127 BY SIMILARITY.  
SQ SEQUENCE 358 AA; 41834 MW; 88344321F24B77C6 CRC64;  
-----

Query Match 63.6%; Score 1157.5; DB 1; Length 358;  
Best Local Similarity 64.7%; Pred. No. 1.3e-71;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:02:46 ; Search time 17.01 Seconds  
(without alignments)  
747.954 Million cell updates/sec

Title: US-09-671-050-10  
Perfect score: 1820  
Sequence: 1 MEKYEKLAKTGEGSYGVFK.....RKARNEGRRRQOVLPLKS 347

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues  
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1157.5	63.6	358	1 KKIA_HUMAN	Q00532 homo sapien
2	694.5	38.2	1030	1 STK9_HUMAN	Q07639 homo sapien
3	565	31.0	305	1 CDK3_HUMAN	Q00526 homo sapien
4	563	30.9	302	1 CD2_CARAU	P51958 carassius a
5	561	30.8	297	1 CDK2_XENLA	P23437 xenopus lae
6	558.5	30.7	292	1 CC2H_DICDI	P34117 dictyosteli
7	554	30.4	298	1 CDK2_HUMAN	P24941 homo sapien
8	551.5	30.3	294	1 CC21_ORYSA	P29618 oryza sativ
9	551.5	30.3	294	1 CC2A_ARATH	P24100 arabidopsis
10	550	30.2	298	1 CDK2_CRIGR	O55076 cricetus
11	548	30.1	298	1 CDK2_RAT	Q63699 rattus norv
12	546	30.0	298	1 CDK2_MESAU	P48963 mesocricetu
13	545	29.9	297	1 CC2_HUMAN	P06493 homo sapien
14	545	29.9	298	1 CDK2_CARAU	P43450 carassius a
15	542	29.8	297	1 CC2_BOVIN	P48734 bos taurus
16	541.5	29.8	294	1 CC2_MAIZE	P23111 zea mays (m
17	541	29.7	297	1 CC2_RAT	P39951 rattus norv
18	540.5	29.7	297	1 CC2_DROME	P23572 drosophila
19	540.5	29.7	302	1 CC22_XENLA	P24033 xenopus lae
20	540	29.7	302	1 CC21_XENLA	P35667 xenopus lae
21	539	29.6	303	1 CC2_CHICK	P13863 gallus gall
22	537	29.5	288	1 CC2H_PLAFK	Q07785 plasmodium
23	536.5	29.5	294	1 CC2_VIGUN	P52389 vigna ungui
24	536.5	29.5	311	1 CC23_TRYBB	P54666 trypanosoma
25	532	29.2	297	1 CC2_MOUSE	P11440 mus musculu
26	532	29.2	301	1 CC21_TRYBB	P38973 trypanosoma
27	531	29.2	292	1 CDK5_BOVIN	Q02399 bos taurus
28	531	29.2	292	1 CDK5_RAT	Q03114 rattus norv
29	530.5	29.1	294	1 CC22_MEDSA	Q05006 medicago sa
30	529	29.1	292	1 CDK5_MOUSE	P49615 mus musculu
31	527	29.0	301	1 CC21_TRYCO	P54664 trypanosoma
32	526	28.9	292	1 CDK5_HUMAN	Q00535 homo sapien
33	526	28.9	296	1 CC2_DICDI	P34112 dictyosteli

Matches 211; Conservative 41; Mismatches 41; Indels 33; Gaps 2;

```
QY 1 MEKTEKIAGTGGSGVGVFKCRNKTSGQVAVVKKFVESEDDPVVKKIALREIMLKOLKH 60
Db 1 MEKTEKIAGTGGSGVGVFKCRNKTSGQVAVVKKFVESEDDPVVKKIALREIMLKOLKH 61
QY 61 PNLVNLLEVFRRKRKRLHVFYCDHTLLNELERNPGVAGVKSVMQTLQALNFCIH 120
Db 62 PNLVNLLEVFRRKRKRLHVFYCDHTLLNELERNPGVAGVKSVMQTLQALNFCIH 121
QY 121 NCHROKIPENILITKQIKICDFGAQILIT-PCDAYTQVATRYRAPPELLVGDFOYG 179
Db 122 NCHROKIPENILITKHSVKIKLDFGFAELLTGFSDYTYDVATRYRSPPELLVGDFOYG 181
QY 180 SSVDIWAICVFAELLTGQPLWPKGSDVDOLYLIRTLVETGFRHVDQAGLELLTSSDPP 239
Db 182 PPDVWAICVFAELLTGQPLWPKGSDVDOLYLIRTKYL----- 219
QY 240 AVASQAGITGKLIPIHQSTFKSNGFFHGISIPEPEDMETLEEKTSVHPVALNFMKGCL 299
Db 220 -----GDLTPRHQQVFSTNOYFSGVKIPDPEDMEPLEKFPNISYPALGKGL 269
QY 300 KMPDRLTCSQLLESSYFDSFOEAQ 325
Db 270 HMDPTERTLCQLLHPHYENIREIE 295
```

## RESULT 2

```
STK9_HUMAN
ID STK9_HUMAN STANDARD; PRT; 1030 AA.
AC 076039; Q14198;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE SERINE/THREONINE-PROTEIN KINASE 9 (EC 2.7.1.37).
GN STK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98389628; PubMed=9721213;
RA Montini E., Andolfi G., Caruso A., Buchner G., Walpole S.M.,
RA Mariani M., Consalez G.G., Trump B., Ballabio A., Franco B.;
RT "Identification and characterization of a novel serine-threonine
RT kinase gene from the Xp22 region.";
RL Genomics 51:427-433(1998).
RN [2]
RP SEQUENCE OF 339-789 FROM N.A.
RA Krause S.W., Rehli M., Kreutz M., Schwarzfischer L., Paulauskis J.D.,
RA Andreesen J.D.;
RT "Differential screening leads to novel genetic markers of monocyte
RT to macrophage maturation";
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
CC -!- ENZYME REGULATION: COULD BE ACTIVATED BY PHOSPHORYLATION ON
CC TYROSINE AND THREONINE. PROBABLY BOTH PHOSPHORYLATIONS ARE
CC REQUIRED FOR ACTIVITY.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG, KIDNEY, PROSTATE,
CC OVARY, PLACENTA, PANCREAS AND TESTIS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDK3 SUBFAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
CC -!- CAUTION: REF 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 415.
```

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```
CC EMBL; Y15057; CA75342.1;
CC EMBL; X89059; CA61445.1; AUT_FRAME.
DR MIM; 300203;
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 13 297 PROTEIN KINASE.
FT NF_BIND 19 27 ATP (BY SIMILARITY).
FT BINDING 42 42 ATP (BY SIMILARITY).
FT ACT_SITE 135 135 BY SIMILARITY.
FT MOD_RES 169 169 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 171 171 (POTENTIAL).
FT MOD_RES 171 171 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT DOMAIN 784 789 POLY-LYS.
FT CONFLICT 339 340 HR -> GT (IN REF. 2).
FT CONFLICT 541 541 L -> W (IN REF. 2).
FT CONFLICT 731 764 MISSING (IN REF. 2).
SQ SEQUENCE 1030 AA; 115537 MW; 8A1C9C438610EF08 CRC64;
```

## Query Match

```
Best Local Similarity 38.2%; Score 694.5; DB 1; Length 1030;
Matches 139; Conservative 69; Mismatches 98; Indels 39; Gaps 5;
QY 1 MEKTEKIAGTGGSGVGVFKCRNKTSGQVAVVKKFVESEDDPVVKKIALREIMLKOLKH 60
Db 10 MKNEELIGVVGAGVGVVLCRKHKETHEIVAIKKKDEENEEVETTELKMLRLTKQ 69
QY 61 PNLVNLLEVFRRKRKRLHVFYCDHTLLNELERNPGVAGVKSVMQTLQALNFCIH 120
Db 70 ENIVELKEAFRRRGRKLYLVEYVEKKNWLEEMPGVPPEKVKSYIYQIKAIHWCHN 129
QY 121 NCHROKIPENILITKQIKICDFGAQILIPGD--AYTDYVATRYRAPPELLVGDFOY 178
Db 130 DIVHRDLKPNLLSHNDVLKDFGFAFNLSGNNANYTEYVATRYRSPPELLLG-APY 188
QY 179 GSSVDIWAICVFAELLTGQPLWPKGSDVDOLYLIRTLVETGFRHVDQAGLELLTSSDP 238
Db 189 GKSVDMWVSGCIGELSDGQPLFPGSEIDQLFTIQKVL----- 227
QY 239 PAVASQAGITGKLIPIHQSTFKSNGFFHGISIPEPEDMETLEEKTSVHPVALNFMKG 297
Db 228 -----GDLTPRHQQVFSTNOYFSGVKIPDPEDMEPLEKFPNISYPALGKGL 269
QY 298 CLKMNPDRLTCSQLLESSYFDSFOEAQIKRKARNEGRNRKQOV 342
Db 277 LKLDLPADRYLLEQCLN---HPTFOTQRLDRSPSPSRAKRPYHV 318
```

## RESULT 3

```
CDK3_HUMAN
ID CDK3_HUMAN STANDARD; PRT; 305 AA.
AC Q00526;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 3 (EC 2.7.1.-).
GN CDK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```



TISSUE=Fetal brain;	
MEDLINE=92347325; PubMed=1639063;	
Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gorka C., Nelson C.,	
Harlow E., Tsai L.-H.;	
"A family of human cdc2-related protein kinases.";	
EMBO J. 11:2909-2917(1992).	
-!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.	
INTERACTS WITH A YET UNKNOWN TYPE OF CYCLIN. CAN PHOSPHORYLATE	
HISTONE H1.	
-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.	
CDK2/CDKX SUBFAMILY.	
-----	
This SWISS-PROT entry is copyright. It is produced through a collaboration	
between the Swiss Institute of Bioinformatics and the EMBL outstation -	
the European Bioinformatics Institute. There are no restrictions on its	
use by non-profit institutions as long as its content is in no way	
modified and this statement is not removed. Usage by and for commercial	
entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
-----	
EMBL; X66357; CAA47001.1; -.	
PIR; S22743; S22743.	
PIR; S23382; S23382.	
HSSP; P24941; IAQL.	
MIM; 123828; -.	
InterPro; IPR00719; Euk_pkinase.	
InterPro; IPR002290; Ser_thr_kin_actsite.	
Pfam; PF00069; pkinase; 1.	
SMART; SM00220; S_TKc; 1.	
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.	
Transferase; Serine/threonine-protein kinase; ATP-binding;	
Cell cycle; Cell division; Mitosis; Phosphorylation.	
DOMAIN 4 286 PROTEIN KINASE.	
NP_BIND 10 18 ATP (BY SIMILARITY).	
BINDING 33 33 ATP (BY SIMILARITY).	
ACT_SITE 127 127 BY SIMILARITY.	
MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).	
MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).	
MOD_RES 160 160 PHOSPHORYLATION (BY SIMILARITY).	
SEQUENCE 305 AA; 35045 MW; 1128BE0096AE262A CRC64;	
-----	
Query Match 31.0%; Score 565; DB 1; Length 305;	
Best Local Similarity 36.8%; Pred. No. 1.3e-31;	
Matches 124; Conservative 66; Mismatches 99; Indels 48; Gaps 9;	
Qy 1 MEKYEKLAKTGESYGVVFKCRNKTSGOVAVKKF---VESEDDPVVKKIALREIRMLKQ 57	
Db 1 MDMFKVEKIGEGTYGVYKAKNRETGQVALKKIRLDEMEGVF---STAIRREISLLKE 57	
Qy 58 LKHPNLNMLEVFRFRKMHVFEYCDHTLLNLELRNPNP-VADGVIKSVLWOTLOALNF 116	
Db 58 LKHPNIVRLDDVHNERKLYLVFEFLSQDLKLYMDSTPGSELPLHLIKSVLFOQQGVSF 117	
Qy 117 CHHNCIHRDKPENILITKQIKCDGFAQIL-IPGDATDYVATRWVAPPELLVGD 175	
Db 118 CHSHRVHRDLKPNLLINELGAIKLAFLADFLGARFGVPLRTTHEVTLWRAPEILLGS 177	
Qy 176 TQYSGSDVMAIGCVFAELLTQPLWPKGSDVDQLYLIIRTLVETGFRHVDQAGLELLTS 235	
Db 178 KFTVTTAVDINSIGIFAEMVTRKALFPDGEIDQLFRPRL----- 219	
Qy 236 SDPPAVASQSA--GITGKLIPRHOSIFKSNFGFHGISIPEPE-DMETLEKFSDDVHPVAL 292	
Db 220 ----GTPSEDTPWPGVT--QLPDYKGSF-----PKWTRKGLLEIVNLEPEGR 260	
Qy 293 NFMKGLKKNPDDRLATCSQLLESSYFDSFQEAQIKRK 329	
Db 261 DLLMQLLQYDPSQRITAKTALAHFVSSPEPSAARQ 297	

RESULT 4				
ID	CC2_CARAU	STANDARD;	PRT;	302 AA.
AC	P51958;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	CELL DIVISION CONTROL PROTEIN 2 HOMOLOG (EC 2.7.1.-) (P34 PROTEIN KINASE) (CYCLIN-DEPENDENT KINASE 1) (CDK1).			
GN	CDK2.			
OS	Carassius auratus (Goldfish).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;			
OC	Cypriniformes; Cyprinidae; Cyprininae; Carassius.			
OX	NCBI_TaxID=7957;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Oocyte;			
RA	Kajiura H., Yamashita M., Katsu Y., Nagahama Y.;			
RT	"Isolation and characterization of goldfish cdc2, a catalytic component of maturation-promoting factor.";			
RL	Dev. Growth Differ. 35:647-654(1993).			
CC	-!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE AND MITOSIS. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE REPLICATIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II. CATALYTIC COMPONENT OF MPF.			
CC	-!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY SIMILARITY).			
CC	-!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH CYCLIN B IN MATURE OOCYTES.			
CC	-!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).			
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	CDK2/CDKX SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
EMBL; D17758; BAA04605.1; -.				
HSSP; P24941; LHCK.				
InterPro; IPR00719; Euk_pkinase.				
InterPro; IPR002290; Ser_thr_kin_actsite.				
Pfam; PF00069; pkinase; 1.				
SMART; SM00220; S_TKc; 1.				
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.				
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.				
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.				
KW	Transferase; Serine/threonine-protein kinase; ATP-binding;			
KW	Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.			
FT	DOMAIN 4 287 PROTEIN KINASE.			
FT	NP_BIND 10 18 ATP (BY SIMILARITY).			
FT	BINDING 33 33 ATP (BY SIMILARITY).			
FT	ACT_SITE 128 128 BY SIMILARITY.			
FT	MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).			
FT	MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).			
FT	MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).			
SQ	SEQUENCE 302 AA; 34499 MW; 58DB812E19B311F5 CRC64;			
-----				
Query Match 30.9%; Score 563; DB 1; Length 302;				
Best Local Similarity 41.2%; Pred. No. 1.8e-31;				
Matches 117; Conservative 60; Mismatches 71; Indels 36; Gaps 6;				
Qy	1 MEKYEKLAKTGESYGVVFKCRNKTSGOVAVKKF-VESEDDPVVKKIALREIRMLKQ 59			
Db	1 MDDYLKIEKIGEGTYGVYKGRNKTGQVAMKKIRLESEEG-VPSTAVREISLLKEIQ 59			

QY 60 HNLVNLIEVFRKRMHLVFEYCDHTLLNELRPNG--VADGVKSVLMQTLQALNFC 117  
 DB 60 HNVVRLDVLVFEFLSDMLKKYLDSPSGQFMDPLMKVSLYQILEGILFC 119  
 QY 118 HHNCIHRDIKPNILITKQIIKICDFGFAQIL-IPGDAYTDVATRWYRAPELLVGD 176  
 DB 120 HCRVRLHDLKPNLLIDNKGVIKLAFLGDLARAFGVVYVTHEVVTLMYRAPEVLGAS 179  
 QY 177 OYSSVDIWAIGCVFAELLTQGPLWPKGSDVQLYLIIRTL----- 217  
 DB 180 RYSPVDVDSIGTFAELATKPLFHDSEIDQLFRITLGTTPNNEVMPDVESLPDYN 239  
 QY 218 -----VETGFRHVDQAGLELLTSS---DPPAVASQASAGIT 249  
 DB 240 TFPKWSGNLSTVKNLKDNGIDLLTKMLIYDPPKRIARQAMT 283

RESULT 5  
 ID CDK2\_XENLA STANDARD; PRT: 297 AA.  
 AC P23437;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-) (CDC2 HOMOLOG EGI PROTEIN KINASE).  
 GN EGI.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-Ovary;  
 RC MEDLINE=91126051; PubMed=1704128;  
 RA Paris J., le Guellec R., Couturier A., le Guellec K., Omilli F.,  
 RA Canonis J., Macneill S., Philippe M.;  
 RT "Cloning by differential screening of a Xenopus cDNA coding for a  
 RT protein highly homologous to cdc2";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:1039-1043(1991).  
 RN [2]  
 RP PHOSPHORYLATION AT THR-160.  
 RX MEDLINE=93345457; PubMed=833783;  
 RA Poon R.Y.C., Yamashita K., Adamczewski J.P., Hunt T., Shuttleworth J.;  
 RT "The cdc2-related protein p40MO15 is the catalytic subunit of a  
 RL protein kinase that can activate p33cdc2 and p34cdc2";  
 RL EMBO J. 12:3123-3132(1993).  
 CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.  
 CC INTERACTS WITH CYCLINS A, D, OR E.  
 CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES  
 CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY  
 CC SIMILARITY).  
 CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED IN UNFERTILIZED EGG, BUT NO  
 CC LONGER MADE IN THE EARLY EMBRYO.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC CDC2/CDKX SUBFAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; X14227; CAA32443.1;  
 DR PIR; A37871; A37871.  
 DR HSSP; P24941; IHCK.  
 DR InterPro; IPR000719; Euk.pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_kin\_actsite.  
 DR Pfam; PF00069; pkinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE\_DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW Cell cycle; Cell division; Mitosis; Phosphorylation.  
 FT DOMAIN 4 286 PROTEIN KINASE.  
 FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
 FT BINDING 33 33 ATP (BY SIMILARITY).  
 FT ACT\_SITE 127 127 BY SIMILARITY.  
 FT MOD\_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 160 160 PHOSPHORYLATION (BY CAK).  
 SQ SEQUENCE 297 AA; 33969 MW; EC30204FCB8D198C CRC64;

Query Match 30.8%; Score 561; DB 1; Length 297;  
 Best Local Similarity 38.2%; Pred. No. 2.4e-31;  
 Matches 116; Conservative 59; Mismatches 89; Indels 40; Gaps 6;

QY 1 MEYKELAKTGGSGVGVFKCRNKTSGOVAVKXFESEDDPVKKIALREIRMLKOLKH 60  
 DB 1 MENFOXVEKIGEGYGVYKARNRETEIVALKKIRLDTEGVPSTAIRSILKELNH 60  
 QY 61 PNLVNLIEVFRKRMHLVFEYCDHTLLNELER-NPNGVADGVKSVLMQTLQALNFC 119  
 DB 61 PNTVKLLDVIHTENKLYLVEFLNQDLAKKMDRSNIGSLSALVKSYLFQLQGLAFCHS 120  
 QY 120 HNCIHRDIKPNILITKQIIKICDFGFAQIL-IPGDAYTDVATRWYRAPELLVGD 178  
 DB 121 HRYLHDLKPNLLIDNKGVIKLAFLGDLARAFGVVYVTHEVVTLMYRAPELLGCKFY 180  
 QY 179 GSSVDIWAIGCVFAELLTQGPLWPKGSDVQLYLIIRTL----- 217  
 DB 181 STAVDWSLGCIFAEMITRAFLPGDSEIDQLFRITLGTTPDEVSWPVTMPDYKSTF 240  
 QY 218 ---VETGFRH---DQAGLELLTSSDPPAVASQASAGITGKLIHQHSIFKNGFPHGS 270  
 DB 241 PKWIRQDSKVVPPLDEGDRLLAOM---LQYDSNKRISAKVALTHP-----FFRDVS 290  
 QY 271 IPEP 274  
 DB 291 RPTP 294

RESULT 6  
 CC2H\_DICDI  
 ID CC2H\_DICDI STANDARD; PRT: 292 AA.  
 AC P34117;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CDC2-LIKE SERINE/THREONINE-PROTEIN KINASE CRP (EC 2.7.1.-).  
 GN CRPA OR CRP.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
 OC NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94032415; PubMed=8218353;  
 RA Michaelis C.E., Weeks G.;  
 RT "The isolation from a unicellular organism, Dictyostelium discoideum,  
 RT of a highly-related cdc2 gene with characteristics of the PCTAIRE  
 RT subfamily";  
 RL Biochim. Biophys. Acta 1179:117-124(1993).  
 CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES  
 CC THE ENZYME (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC CDC2/CDKX SUBFAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; X14227; CAA32443.1;  
 DR PIR; A37871; A37871.  
 DR HSSP; P24941; IHCK.  
 DR InterPro; IPR000719; Euk.pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_kin\_actsite.  
 DR Pfam; PF00069; pkinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.

of a cdc28 mutation in *Saccharomyces cerevisiae*, is a homolog of  
Xenopus Egl.  
EMBO J. 10:2653-2659(1991).

[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-91367262; PubMed-1653904;  
RA Tsai L.-H., Harlow E., Meyerson M.;  
RT "Isolation of the human cdk2 gene that encodes the cyclin A- and  
adenovirus E1A-associated p33 kinase."  
RL Nature 353:174-177(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92020980; PubMed-1717994;  
RA Ninomiya-Tsuji J., Nomoto S., Yasuda H., Reed S.L., Matsumoto K.;  
RT "Cloning of a human cDNA encoding a CDC2-related kinase by  
complementation of a budding yeast cdc28 mutation."  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9006-9010(1991).  
RN [4]  
RP PHOSPHORYLATION SITES.  
MEDLINE-93010995; PubMed-1396589;  
RA Gu Y., Rosenblatt J., O'Morgan D.O.;  
RT "Cell cycle regulation of CDK2 activity by phosphorylation of Thr160  
and Tyr15."  
RL EMBO J. 11:3995-4005(1992).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
MEDLINE-93288132; PubMed-8510751;  
RA de Bondt H.L., Rosenblatt J., Jancarik J., Jones H.D.,  
Morgan D.O., Kim S.-H.;  
RT "Crystal structure of cyclin-dependent kinase 2."  
RL Nature 363:595-602(1993).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CYCLIN A.  
MEDLINE-95356811; PubMed-7630397;  
RA Jeffrey P.D., Russo A.A., Polyak K., Gibbs E., Hurwitz J.,  
Massague J., Pavletich N.P.;  
RT "Mechanism of CDK activation revealed by the structure of a  
cyclinA-CDK2 complex."  
RL Nature 376:313-320(1995).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.33 ANGSTROMS) OF COMPLEX WITH L868276.  
MEDLINE-96181476; PubMed-8610110;  
RA de Azevedo W.F. Jr., Muleer-Diekmann H.-J., Schulze-Gahmen U.,  
Worland P.J., Sauville E., Kim S.-H.;  
RT "Structural basis for specificity and potency of a flavonoid  
inhibitor of human CDK2, a cell cycle kinase."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:2735-2740(1996).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH CG2A AND KIP1.  
MEDLINE-96300318; PubMed-8684460;  
RA Russo A.A., Jeffrey P.D., Patten A.K., Massague J., Pavletich N.P.;  
RT "Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor  
bound to the cyclin A-CDK2 complex."  
RL Nature 382:325-331(1996).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH CG2A.  
MEDLINE-96313126; PubMed-8756328;  
RA Russo A.A., Jeffrey P.D., Pavletich N.P.;  
RT "Structural basis of cyclin-dependent kinase activation by  
phosphorylation."  
RL Nat. Struct. Biol. 3:696-700(1996).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
MEDLINE-97075215; PubMed-8917641;  
RA Schulze-Gahmen U., de Bondt H.L., Kim S.-H.;  
RT "High-resolution crystal structures of human cyclin-dependent kinase  
2 with and without Arp: bound waters and natural ligand as guides for  
inhibitor design."  
RL J. Med. Chem. 39:4540-4546(1996).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
MEDLINE-97475219; PubMed-9334743;  
RA Lawrie A.M., Noble M.E.M., Tunnah P., Brown N.R., Johnson L.N.,

modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
CC EMBL; L00652; AAA16056.1; .  
DR PIR; S40021; S40021.  
DR HSSP; P24941; 1A01.  
DR DictyDb; DD05039; crpA.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_kin\_actsite.  
DR Pfam; PF00069; pkinase; 1.  
DR SMART; SM00220; S\_TK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR Transferase; Serine/threonine-protein kinase; ATP-binding;  
KW Phosphorylation.  
FT DOMAIN 4 285 PROTEIN KINASE.  
FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
FT BINDING 33 33 ATP (BY SIMILARITY).  
FT ACT\_SITE 126 126 BY SIMILARITY.  
FT MOD\_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 292 AA; 361AB54C4E3BD41E CRC64;

Query Match 30.7%; Score 558.5; DB 1; Length 292;  
Best Local Similarity 35.0%; Pred. No. 3.4e-31;  
Matches 115; Conservative 66; Mismatches 101; Indels 47; Gaps 6;  
QY 1 MEKYELAKTEGSGYGVFKCRNKTSGQVAVKFESEDDPVKKIALREIRMLKOLKH 60  
DB 1 MEKYSIEKLEGEGYGVNKKRKNRGEIVALKRILDSDEGVCPTAIRLSILKELKH 60  
QY 61 PNLVNLIEVFRKKMHLVPEYCDHTLNELELRNPNGVADGVKSVLWQTLQALNFCCHI 120  
DB 61 PNLVRLDVIHTERKLTLPVEYLDQKLKYLDECGGEISPTIKSFMYOLKGVAFCHDH 120  
QY 121 NCHROIKENILITKOGIILKICDGFQAOL-IPGDAYTDYVATRWYRAPELLVGDQYQ 179  
DB 121 RVLRDLKPNLNLNKRGEIKLADFLARAFGIPVRYSHVETLWYRAPDVLGSRKYS 180  
QY 180 SSVDIWAICVFAELLTGQPLWPKGSDVDQLYLIRLVTFGRHVDQAGLELLTSSDPP 239  
DB 181 TPDIWSALCIFAEMASGRPLFGSGTSDQLFRIFKILGTP-----NESWP 227  
QY 240 AVASQASGITKLIPIRHOSIF-----KNGFFHGISIPEDMETLEEFSDVHPVALN 293  
DB 228 SITE-----LPEYKDFVHPAQLSSIVHG-----LDEK-----GLN 260  
QY 294 FMKGLKMPDRLTCSQLESSEYFDSFQ 322  
DB 261 LLSKMLQYDPNQRTAAALAKHPYFDGLE 289

RESULT 7  
ID CDK2\_HUMAN STANDARD; PRT; 298 AA.  
AC P24941;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE CELL DIVISION PROTEIN KINASE 2 (BC 2.7.1-) (P33 PROTEIN KINASE).  
GN CDK2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-91330891; PubMed-1714386;  
RA Elledge S.J., Spottswood M.R.;  
RT "A new human p34 protein kinase, CDK2, identified by complementation





Db 61 SNIVKQDVVHSEKRLYLVEYLDLKKHMDSTPDFSKOLHMIKTYLQILRGIAVCHS 120  
 QY 120 HNCIHRDKPENILTKO-GIHKICDFGFAQIL-IPGDAYDVVAVRWRAPELLVGDQ 177  
 Db 121 HRVLRDLAPQNLDRRTNSLKLADGLARAFGIPVTFTHVTLWYRAPELLGSHH 180  
 QY 178 YGSSVDIWAIGCVFAELLTGOPWPKSDVQOLYLIR- - - - -TLVETGFRHVDQAGLELLT 234  
 Db 181 YSTPVDINSVGVFAEMISOKPLFGDSEIDQLKFRIMGTPIEDTWK- - - - - 229  
 QY 235 SSDDPAVASQSAGTGLIPRHOSIFKSNFGFHGISIPEPEDEMETLEEKFSDVHPVALNF 294  
 Db 230 - - - - -GVTS--LPDYKSAFPK- - - - -WKPTDLETF- - - - -VNLDPDGVDL 263  
 QY 295 MKGCLKKNPDDRLTCSQLLESSYF 318  
 Db 264 LSKMLLMDPTKRINARAALHEYF 287

## RESULT 10

CDK2\_CRIGR  
 AC 05076; STANDARD; PRT; 298 AA.  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.1-).  
 GN CDK2.  
 OS Cricetus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus.  
 OX NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ellenrieder C., Bartosch B., Lee G.Y., Murphy M., Sweeney C.,  
 RA Hergersberg M., Hunt T., Carrington M., Jaussi R.;  
 RT "The 39 kDa form of CDK2 arises through alternative splicing, is  
 RT expressed in many but not all mammals, and is an active kinase.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.  
 CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL  
 CC DURING S PHASE AND G2.  
 CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES  
 CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY  
 CC SIMILARITY).  
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CDK2-ALPHA (SHOWN HERE) AND  
 CC CDK2-BETA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC CDC2/CDKX SUBFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: AJ223949; CAAL1680.1; -;  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase: Serine/threonine-protein kinase; ATP-binding; Cell cycle;  
 KW Cell division; Mitosis; Phosphorylation; Alternative splicing.  
 FT DOMAIN 4 286  
 FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
 FT BINDING 33 33 ATP (BY SIMILARITY).  
 FT

FT ACT\_SITE 127 127 BY SIMILARITY.  
 FT MOD\_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 160 160 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).  
 SQ SEQUENCE 298 AA; 33873 MW; 3B580DC2460CD8E CRC64;

Query Match 30.2%; Score 550; DB 1; Length 298;  
 Best Local Similarity 38.1%; Pred. No. 1.3e-30;

Matches 117; Conservative 55; Mismatches 89; Indels 46; Gaps 6;

QY 1 MEKYELAKTGEISYGVVFKCRNKTSGQVAVKRFVESEDDPVPKKIALREIRMLQALKH 60  
 Db 1 MENFOVKEKIGEGTYGVVYKAKNKUTGEVALKKIRLDITETGVSFSTAREISLLKELNH 60  
 QY 61 PNLVNLIEVFRKHKMHLVEYCDHTLLNELRNP-NGVADGVTKSVLMOTLAALNFCHI 119  
 Db 61 PNLVNLIEVFRKHKMHLVEYCDHTLLNELRNP-NGVADGVTKSVLMOTLAALNFCHI 119  
 QY 120 HNCIHRDKPENILTKO-GIHKICDFGFAQIL-IPGDAYDVVAVRWRAPELLVGDQ 178  
 Db 121 HRVLRDLAPQNLDRRTNSLKLADGLARAFGIPVTFTHVTLWYRAPELLGCKYY 180  
 QY 179 GSSVDIWAIGCVFAELLTGOPWPKSDVQOLYLIR- - - - - 217  
 Db 181 STAVDINSVGVFAEMISOKPLFGDSEIDQLKFRIMGTPIEDTWK- - - - - 240  
 QY 218 - - - - -VETGFRHV- - - - -DOAGLELLTSS- - - - -DPPAVASQSAGTGLIPRHOSIFKSNFGFH 267  
 Db 241 PKWARQDFSKVVPPLDEGSRLLSQMLHYDPNKRISAKAAL- - - - -AHPFFQ 287  
 QY 268 GISIPEP 274  
 Db 288 DVTKEVP 294

## RESULT 11

CDK2\_RAT  
 ID CDK2\_RAT STANDARD; PRT; 298 AA.  
 AC Q63699; O09136;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.1-).  
 GN CDK2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thyroid;  
 RX MEDLINE=95166553; PubMed=7862443;  
 RA Kotani S., Endo T., Kitagawa M., Higashi H., Onaya T.;  
 RT "A variant form of cyclin-dependent kinase 2 (Cdk2) in a malignantly  
 RT transformed rat thyroid (FRTL-Tc) cell line.";  
 RL Oncogene 10:663-669(1995).  
 RN [2]  
 RP SEQUENCE OF 19-124 FROM N.A.  
 RX MEDLINE=96113578; PubMed=8673024;  
 RA Hosokawa Y., Yang M., Kaneko S., Tanaka M., Nakashima K.;  
 RT "Synergistic gene expressions of cyclin E, cdk2, cdk5 and E2F-1  
 RT during the prolactin-induced G1/S transition in rat Nb2 pre-T  
 RT lymphoma cells.";  
 RL Biochem. Mol. Biol. Int. 37:393-399(1995).  
 CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.  
 CC INTERACTS WITH CYCLINS A, D, OR E. ACTIVITY OF CDK2 IS MAXIMAL  
 CC DURING S PHASE AND G2.  
 CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES  
 CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY  
 CC SIMILARITY).  
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CDK2-ALPHA (SHOWN HERE) AND  
 CC CDK2-BETA; ARE PRODUCED BY ALTERNATIVE SPLICING.



```

QY 268 GISPEP 274
Db 288 DVTKPVP 294

RESULT 13
CC2_HUMAN
ID CC2_HUMAN STANDARD; PRT; 297 AA.
AC P06493;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG (EC 2.7.1.-) (P34 PROTEIN
DE KINASE) (CYCLIN-DEPENDENT KINASE 1) (CDK1).
GN CDC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RS SEQUENCE FROM N.A.
RX MEDLINE=87201915; PubMed=3553962;
RA Lee M.G., Nurse P.;
RE "Complementation used to clone a human homologue of the fission yeast
RE cell cycle control gene cdc2.";
RL Nature 327:31-35(1987).
RN 12;
RS PHOSPHORYLATION, AND ASSOCIATION WITH P13.
RX MEDLINE=89253421; PubMed=3289755;
RA Draetta G., Beach D.;
RE "Activation of cdc2 protein kinase during mitosis in human cells;
RE cell cycle-dependent phosphorylation and subunit rearrangement.";
RL Cell 54:17-26(1988).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE
CC AND MITOSIS. P34 IS A COMPONENT OF THE KINASE COMPLEX THAT
CC PHOSPHORYLATES THE REPETITIVE CARBOXYL-TERMINUS OF RNA
CC POLYMERASE II.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME. WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT.
CC -!- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH A REGULATORY
CC SUBUNIT AND WITH A CYCLIN.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05360; CAA28963.1; -
DR EMBL; Y00272; CAA68376.1; -
DR PIR; A29539; A29539.
DR HSP; P24941; 1HCK.
DR SWISS-2DPAGE; P06493; HUMAN.
DR MM; 116940; -
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation.
FT DOMAIN 4 287
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).

```

---

```

FT ACT_SITE 128 128 BY SIMILARITY.
FT MOD_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).
SQ SEQUENCE 297 AA; 34095 MW; 942D79448E490A CRC64;

Query Match 29.9%; Score 545; DB 1; Length 297;
Best Local Similarity 40.2%; Pred. No. 2.9e-30;
Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;

QY 1 MKYKELKAGTGGSGYGVFKCNKTSQGVAVKKF-VESEDDPVVKKIALREIRMLKOLK 59
Db 1 MEDYTKETKIGEGTVGVYKGRHKTGQVAMKKIRLESEEG-VPSTAREISLLEKLR 59
QY 60 HENLVNLTIEVFRKRKMLVFEYCDHTLLNELEPNPG--VADGVKSVLWQTLQALNFC 117
Db 60 HENIVSLQDVLQMDSRLYLIFEFLSMDLKKYLDSTPPQYMDSSLVKSXYLIQLOGIVFC 119
QY 118 HTNHCIRHDKPENILITKQIILKICDFGFAQL-IPGDVDTDYVATRWYRAPELLVGDT 176
Db 120 HSRVLRHDLKPNQLIDDKTGLADFGIARAFGIPIRVYTHEVVTLMWYRSPVLLGSA 179
QY 177 QYGSVDIMAIGCVFAELITGQPLWPKGSDVQLYLIIRTL----- 217
Db 180 RYTPVDVINSIGTIFAEELATKKPLFHGDSIDLQIFRIFALGTPNNEVWPEVESLDQYKN 239
QY 218 -----VETGFRHVDQAGLELLTSS---DPPAVASQSAGITGKLIPIRH 256
Db 240 TFPKWKPGSLASHVKNLQENGLDLSLMLTYDP-----AKRISGKMALNH 284

RESULT 14
ID CDK2_CARAU STANDARD; PRT; 298 AA.
AC P43450;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION PROTEIN KINASE 2 (EC 2.7.1.-).
GN CDK2.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
OX NCBI_TaxID=7957;
RN 1;
RS SEQUENCE FROM N.A.
RX TISSUE=Oocyte;
RX MEDLINE=92333802; PubMed=1339336;
RA Hirai T., Yamashita M., Yoshikuni M., Tokumoto T., Kajiuura H.,
RA Sakai N., Nagahama Y.;
RE "Isolation and characterization of goldfish cdk2, a cognate variant
RE of the cell cycle regulator cdc2.";
RL Dev. Biol. 152:113-120(1992).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.
CC INTERACTS WITH CYCLINS A, D, OR E.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-160 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S40289; AAB22550.1; -
DR PIR; A44878; A44878.

```





Fri May 3 11:22:25 2002

us-09-671-050-10.rsp

Page 12

. . . .

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 11:02:21 ; Search time 40.57 Seconds  
(without alignments)  
1251.084 Million cell updates/sec

Title: US-09-671-050-10  
Perfect score: 1820  
Sequence: 1 MEXYKLAKTGGSGYGVFK.....RKARNEGRRRQOVLPLKS 347

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues 473505  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	1001.5	55.0	353	5	Q9U2H1		Q9u2h1 caenorhabdi
2	1001.5	55.0	392	5	Q9VMN3		Q9vmn3 drosophila
3	941.5	51.7	566	6	Q9PTK0		Q9ptk0 oryctolagus
4	915.5	50.3	493	4	Q92772		Q92772 homo sapien
5	904.5	49.7	564	11	Q9QYI2		Q9qyi2 mus musculus
6	904.5	49.7	568	11	Q9QUK0		Q9quk0 mus musculus
7	902	49.6	329	11	Q9QYI1		Q9qyi1 mus musculus
8	758.5	41.7	455	4	Q9P1L4		Q9p1l4 homo sapien
9	755	41.5	657	5	Q9BMG2		Q9bmg2 trypanosoma
10	743.5	40.9	1106	5	Q9GRT9		Q9grt9 leishmania
11	737.5	40.5	457	11	Q9JM02		Q9jmo2 rattus norv
12	737.5	40.5	505	11	Q9TM01		Q9tmo1 rattus norv
13	618	34.0	1104	13	Q9W6R6		Q9w6r6 fuigu rubrip
14	615.5	33.8	997	4	Q9QUL6		Q9qul6 homo sapien
15	565.5	31.1	294	10	Q9ZRL1		Q9zrl1 tritium ae
16	551	30.3	288	5	Q96821		Q96821 plasmodium
17	551	30.3	288	5	Q9XZD6		Q9xzd6 plasmodium
18	549.5	30.2	303	13	Q9DGA2		Q9dga2 oryzias jav
19	546.5	30.0	294	10	Q9M307		Q9m307 arabidopsis

20	546.5	30.0	303	13	Q9DGA1	Q9dga1 oryzias jav
21	545.5	30.0	294	10	Q82666	Q82666 brassica na
22	543.5	29.9	303	13	Q9DGD3	Q9dgd3 oryzias lat
23	542.5	29.8	294	10	P93101	P93101 chenopodium
24	542.5	29.8	300	5	O15890	O15890 toxoplasma
25	541.5	29.8	297	5	Q9TX69	Q9tx69 drosophila
26	541.5	29.8	303	13	Q9DGA5	Q9dga5 oryzias cur
27	541.5	29.8	303	13	Q9DGP8	Q9dgp8 oryzias luz
28	540.5	29.7	297	5	Q9TX74	Q9tx74 drosophila
29	539.5	29.6	294	10	O40750	O40750 pinus conto
30	539.5	29.6	297	5	Q9TX73	Q9tx73 drosophila
31	539.5	29.6	300	5	O44000	O44000 toxoplasma
32	539.5	29.6	300	5	O17066	O17066 asterina pe
33	539	29.6	298	5	O27032	O27032 theileria p
34	538.5	29.6	294	10	O40789	O40789 petroselinu
35	538.5	29.6	294	10	O82135	O82135 pisum sativ
36	538.5	29.6	300	3	O13379	O13379 pneumocysti
37	538.5	29.6	300	3	O13380	O13380 pneumocysti
38	537.5	29.5	297	5	Q9TX71	Q9tx71 drosophila
39	537	29.5	288	5	O96820	O96820 plasmodium
40	536.5	29.5	297	5	Q9TX68	Q9tx68 drosophila
41	535.5	29.4	294	10	P93556	P93556 sesbania ro
42	535.5	29.4	294	10	O41639	O41639 vigna aconi
43	535.5	29.4	294	10	O43361	O43361 picea abies
44	534.5	29.4	294	10	O40484	O40484 nicotiana t
45	534.5	29.4	294	10	Q9FUR4	Q9fur4 nicotiana t

ALIGNMENTS

RESULT 1  
Q9U2H1 PRELIMINARY: PRT; 353 AA.  
AC Q9U2H1:  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Y42A5A.4 PROTEIN.  
GN Y42A5A.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;  
OC Rhabditiidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Steward C.A.;  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RX MEDLINE=9069613; PubMed=9851916;  
RA none;  
RT \*Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology.\*;  
RL Science 282:2012-2018(1998).  
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AL032618; CAB63367.1; -.  
DR HSP; P24941; IHCL.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_kin\_actsite.  
DR Pfam; PF00069; pkinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 353 AA; 40665 MW; 7E32260D9849CAC9 CRC64;

Query Match 55.0%; Score 1001.5; DB 5; Length 353;  
Best Local Similarity 53.4%; Pred. No. 4.6e-80;  
Matches 187; Conservative 59; Mismatches 63; Indels 41; Gaps 3;

[illegible]

RESULT	2
Q9VWN3	
ID	Q9VWN3
AC	PRELIMINARY;
DT	01-MAY-2000
DT	01-MAY-2000 (TReMBLrel. 13, Created)
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT	01-JUN-2000 (TReMBLrel. 17, Last annotation update)
DE	Cf7238
PRT:	392 AA.

OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RN NCBI\_TaxID=7227;  
 ON f11

SEQUENCE FROM N.A.  
STRAIN-BERNELLEY;  
MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Plamkoch C., Miklos G.L.G.,  
RA Balow R.M., Basu A., Bayendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhendari D., Baldwin D.G.,  
RA Borkova D., Botchan M.R., Bouck J., Brakstein P., Brotier P.,  
RA Burtis J.M., Busam D.A., Butler H., Cadieux E., Center A., Chandra I.,  
RA Cherry J.K., Cawley S., Dahlke S., Davidson L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Lew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.J., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Liao P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshirefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Neilson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese H.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodgate T., Ye J., Yang S., Tao Q.A.,  
RA Ye J., Voh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:  
RT "The genome sequence of *Drosophila melanogaster*.";  
RT Science 287:2185-2195(2000).  
CC 1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC EMBL: AF003611; AAF52279.1; -.  
CC HSSP: P24941; IWC1.  
DR FlyBase: FB050031730; CG7236.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR TaterPro: IPR002290; Ser\_thr\_kin\_actsite.  
DR Pfam: PF00069; pkinase; 1.  
DR SMART: SM00220; S\_TKC; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
DR AIP-binding: Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 392 AA; 44962 MW; C9627D007784E3D CRC64.

Query Match	55.08;	Score 1001.5;	DB 5;	Length 392;
Best Local Similarity	53.09;	Pred. NO. 5.3e-80;		
Matches 184;	Conservative	65;	Mismatches 61;	Indels 37;
			Gaps	3;

```

QY      1 MEKTEKLANTGGSGVWFKCRNKTSGQVAVKFVESEDDPVVRKIALREIRMLKOLKH 60
Db      1 MDRYEKLSRLGEGSYGVYKCDREPTGALVAVKFVESEDDPAIRKIALREIRLLKNLKH 60

QY      61 PNLVNLTVEFRKKRMHLVFECYDHTLNELERNPVGADVTKSVLMQOTLOANFCUHLH 120
Db      61 PNLVSLEVEFRKKRHLVFEECELTVLJHELRHPGCCPEHTKOICFTQLUGVAVCHQK 120

QY      121 NCTHRDIKPENILITKGOIHKIDCFGAQIILIPGDATVYATRWYRAPELLVGDTOYGS 180
Db      121 GCULHRDIKPENILITTAQQWKIDCFGFARMLSPENYTYDVATRWYRAPELLVGDTOYGT 180

QY      181 SVDIWAIGCVFAELLTGCPUMPKGSDVDQLYLIIRTLETGTRHVDQAGLELLTSSDPRA 240
Db      181 PVDVMAIGCLFAELVRGEALWPGRSVDQLYLKTL----- 217

QY      241 VASQSAGITGKLIPIRHOSITKNGFHFGISITPEDEMETLEBK F---SDHVFPVALFMKG 297
Db      218 -----GDLPRHITOIFQGNEYFKGIILIPPTLEEDKMPAKSQOQNPLTIDFLKK 268

QY      298 CLKMPPDDR LTCSSLSESYTFDSF--QBAQIKRKARNENRRNQOV 342
Db      269 CLDKDPTKRWCSEKTHSFYDYIAKORELEHVNSLEANLRQOOL 315

```

```

RESULT      3
Q9TKKO      PRELIMINARY;          PRT;    566 AA.
AC
DI Q9TKKO
DI 01-MAY-2000 (TREMBLrel. 13, Created)
DI 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DI 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SER/THR KINASE KKIAMRE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Amniota; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Gomi H., Sun W., Finch C.E., Itohara S., Yoshimi K., Thompson R.F.;
RT "Learning induces a cdcl2-related protein kinase kkiamre.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDAJ databases.
RC -! SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029045; BAA88508.1; .
DR HSSP; P24941; 1HCL.

```

Fri May 3 11:22:26 2002

```
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00069; pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 566 AA; 64052 MW; 5D5759550902EA9 CRC64;
SQ

Query Match 51.7%; Score 941.5; DB 6; Length 566;
Best Local Similarity 49.7%; Pred. No. 1.7e-74;
Matches 177; Conservative 67; Mismatches 65; Indels 47; Gaps 6;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKOLKH 60
DB 1 MEKENLGLVGECSYGMVKCRNKDGRIVAIKKFLESDDDKMKVKIAMREIKLKLQRH 60
QY 61 PNLVNLIEVFRKRKRMHLVFEYCDHTLLNELERNPNGVADGVKSVLWOTLQALNFCIH 120
DB 61 ENLVNLIEVCKKKRWYLVFEVDHTLLDLELFPNGLDQVQKYLFLQILINGICFCHSH 120
QY 121 NCIIHRDKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDYQYG 179
DB 121 NIHRDKPENILVSQSVGVKLCDFGFARTLAAPGEVTDYVATRWYRAPELLVGDYKYG 180
QY 180 SSVDIWAIGCVFAELLTGQPLWPKGSDVDOLYLIIRTLVETGFRHVDQAGLELLTSSDPP 239
DB 181 KAVDVWAIGCVTEMGMGEPLFPDSDIDQLYIMKCL----- 218
QY 240 AVASQASGITKLIIPRHOSIFKSNPFHGSIPEPEMETLEKFSVDVHPVALNFMKGL 299
DB 240 AVASQASGITKLIIPRHOSIFKSNPFHGSIPEPEMETLEKFSVDVHPVALNFMKGL 299
QY 219 -----GNLIIPRHQELFYNFVAGVRLPEIKESPLERRYPKLSSEVIDLAKKCL 268
QY 300 KWNPDRLTCSQLESSEYF--DSF-----QEAQIK--RKARN-----EGNRNRQO 341
DB 269 HIVDPDRPFCAELLHHDFFQMDGFAERFSQELQKQVQKDARNVSLSKSKSNRKKKEK 324

RESULT 5
Q9QY12 PRELIMINARY; PRT; 564 AA.
ID Q9QY12 AC Q9QY12
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SER/THR KINASE KIAMRE-GAMMA.
GN CDKL2 OR KKM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=BRAIN;
RA Sessa T., Goni H., Sun W., Ikeda T., Thompson R.F., Itohara S.;
RT "The Murine KIAMRE gene: variants, dual promoters, expression and
chromosomal localization.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB029066; BAA88428.1; -.
DR HSP; P27703; 1ERK.
DR MGD; MGI:1858227; Cdkl2.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 564 AA; 63640 MW; DACDEF630CCEB82D5 CRC64;

Query Match 49.7%; Score 904.5; DB 11; Length 564;
Best Local Similarity 48.6%; Pred. No. 3.2e-71;
Matches 174; Conservative 66; Mismatches 69; Indels 49; Gaps 7;

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKOLKH 60
DB 1 MEKENLGLVGECSYGMVKCRNKDGRIVAIKKFLESDDDKMKVKIAMREIKLKLQRH 60

QY 1 MEKYEKLAKTGEGSYGVVFKCRNKTSGQVAVKKEVESEDDPVVKKIALREIRMLKOLKH 60
DB 1 MEKENLGLVGECSYGMVKCRNKDGRIVAIKKFLESDDDKMKVKIAMREIKLKLQRH 60
```

QY 61 PNLVNLIEVFRKRMHVLVEFYCDHTLLNELERNPNVAGVGVKSVLWQTLQALNFCIH 120  
 Db 61 ENLVNLLEVCKKKRWLVFEVDHTILDLLKLPNGLDYQVVKYLFQINGIGFCHSH 120  
 QY 121 NCIHRIKPNILITKOGIIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDYQYG 179  
 Db 121 NIIRHDIKPNILVSQGVVKLCDFGFARTLAAGGEVYTDYVATRWYRAPELLVGDYQYG 180  
 QY 180 SSVDAWAGCVFAELLTGOPLWPGKSDVDQYLYLIRTLVETGFRHVDQAGLELLTSSDPP 239  
 Db 181 KAVDIWAIGCLVIEMLMGQPLPFGESDIDQLHIMTCL----- 218  
 QY 240 AVASQAGITGKLIPIRHQSIFKSNFFHGHSIPEPEDMET--LEEKFSVDHPVALNFMKG 297  
 Db 219 -----GNLIPRHOELFYKNPVFAGVRLPVKDAEAPLESRYPKLPEAVISLAKK 268  
 QY 298 CLKMNPDDRITCSOLLESYF--DSF-----QEAQIK--RKARN-----BGRNRRQO 341  
 Db 269 CLHIDPDRPCADLLRHDFDFOMDGFARFSQELQLKIEKDARNNSLPKKSQRKKEK 326

RESULT 6  
 Q90UK0  
 ID Q90UK0 PRELIMINARY: PRT: 568 AA.  
 AC Q90UK0  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE SER/THR KINASE KTIAMRE. 17, Last annotation update)  
 GN CDK12 OR KKM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ, AND C57BL/6;  
 RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;  
 RT "The Murine KTIAMRE gene: variants, dual promoters, expression and  
 chromosomal localization".  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB029073; BAA88439.1;  
 DR EMBL; AB029068; BAA88439.1; JOINED.  
 DR EMBL; AB029070; BAA88439.1; JOINED.  
 DR EMBL; AB029071; BAA88439.1; JOINED.  
 DR EMBL; AB029072; BAA88439.1; JOINED.  
 DR EMBL; AB029065; BAA88427.1;  
 DR HSP; P27703; IERK.  
 DR MGD; MGI:1858227; Cdk12.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_kin\_actsite.  
 DR Pfam; PF00069; pkinase; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 FT NON\_TER 329 329  
 SQ SEQUENCE 568 AA; 64055 MW; A43B75E2E9EB86C4 CRC64;

Query Match 49.7%; Score 904.5; DB 11; Length 568;  
 Best Local Similarity 48.6%; Pred. No. 3.2e-71;  
 Matches 174; Conservative 66; Mismatches 69; Indels 49; Gaps 7;  
 QY 1 MEKYEKLAKTGEGSYGVFKCRNKTSGQVAVKFFVESEDDPVVKKIALREIRMLKQLKH 60  
 Db 1 MEKYENLGLVGEESYGVNWKCRNKGRIVAIKKFLSEDDDKMKVKIAMREIKLLKQLRH 60  
 QY 61 PNLVNLIEVFRKRMHVLVEFYCDHTLLNELERNPNVAGVGVKSVLWQTLQALNFCIH 120  
 Db 61 ENLVNLLEVCKKKRWLVFEVDHTILDLLKLPNGLDYQVVKYLFQINGIGFCHSH 120

QY 121 NCIHRIKPNILITKOGIIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDYQYG 179  
 Db 121 NIIRHDIKPNILVSQGVVKLCDFGFARTLAAGGEVYTDYVATRWYRAPELLVGDYQYG 180  
 QY 180 SSVDAWAGCVFAELLTGOPLWPGKSDVDQYLYLIRTLVETGFRHVDQAGLELLTSSDPP 239  
 Db 181 KAVDIWAIGCLVIEMLMGQPLPFGESDIDQLHIMTCL----- 218  
 QY 240 AVASQAGITGKLIPIRHQSIFKSNFFHGHSIPEPEDMET--LEEKFSVDHPVALNFMKG 297  
 Db 219 -----GNLIPRHOELFYKNPVFAGVRLPVKDAEAPLESRYPKLPEAVISLAKK 268  
 QY 298 CLKMNPDDRITCSOLLESYF--DSF-----QEAQIK--RKARN-----BGRNRRQO 341  
 Db 269 CLHIDPDRPCADLLRHDFDFOMDGFARFSQELQLKIEKDARNNSLPKKSQRKKEK 326

RESULT 7  
 Q9QYI1  
 ID Q9QYI1 PRELIMINARY: PRT: 329 AA.  
 AC Q9QYI1  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE SER/THR KINASE KTIAMRE-DELTA (FRAGMENT).  
 GN CDK12 OR KKM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=BRAIN;  
 RA Sassa T., Gomi H., Sun W., Ikeda T., Thompson R.F., Itohara S.;  
 RT "The Murine KTIAMRE gene: variants, dual promoters, expression and  
 chromosomal localization".  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB029067; BAA88429.1;  
 DR HSP; P27703; IERK.  
 DR MGD; MGI:1858227; Cdk12.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_kin\_actsite.  
 DR Pfam; PF00069; pkinase; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 FT NON\_TER 329 329  
 SQ SEQUENCE 329 AA; 38023 MW; F72EB8965AE8585 CRC64;

Query Match 49.6%; Score 902; DB 11; Length 329;  
 Best Local Similarity 49.0%; Pred. No. 2.4e-71;  
 Matches 173; Conservative 63; Mismatches 73; Indels 44; Gaps 6;  
 QY 1 MEKYEKLAKTGEGSYGVFKCRNKTSGQVAVKFFVESEDDPVVKKIALREIRMLKQLKH 60  
 Db 1 MEKYENLGLVGEESYGVNWKCRNKGRIVAIKKFLSEDDDKMKVKIAMREIKLLKQLRH 60  
 QY 61 PNLVNLIEVFRKRMHVLVEFYCDHTLLNELERNPNVAGVGVKSVLWQTLQALNFCIH 120  
 Db 61 ENLVNLLEVCKKKRWLVFEVDHTILDLLKLPNGLDYQVVKYLFQINGIGFCHSH 120  
 QY 121 NCIHRIKPNILITKOGIIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGDYQYG 179  
 Db 121 NIIRHDIKPNILVSQGVVKLCDFGFARTLAAGGEVYTDYVATRWYRAPELLVGDYQYG 180  
 QY 180 SSVDAWAGCVFAELLTGOPLWPGKSDVDQYLYLIRTLVETGFRHVDQAGLELLTSSDPP 239  
 Db 181 KAVDIWAIGCLVIEMLMGQPLPFGESDIDQLHIMTCL----- 218

Qy	240	AVAS0AGITGKLI	PHQSIKFSNGFHGISIPBEDMET--LEEKFSVDHPVALNFMKG	297
Db	215	-----GNLIPHQELFKVNPAGVRLPEVKDAEAPLESRYPKLPKEAVISLAKK	268	
Qy	298	CLKMNPPDRILTC	SOLLESSYF--DSF-----QEAQIK--RKARNEGRNRQO	341
Db	269	CLHIDDPKRPFCADLLRHDF	FQMDGFQAEFRFSQELQKTEKDARNNSLPPKRSQK	321
RESULT	8			
ID	Q9P114	PRELIMINARY;	PRT;	455 AA.
AC	Q9P114;			
DT	01-OCT-2000 (Tremblrel. 15, Created)			
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)			
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)			
GN	SERINE-THREONINE PROTEIN KINASE NKTAMRE.			
GN	NKTAMRE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TaxID	9606;			
SEQUENCE	FROM N.A.			
RA	Midmer M., Haq R., Zanke B.W.;			
RA	"NKTAMRE a novel kinase deleted in human leukemia."			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	EMBL: AF130372; AAF36509.1; -.			
DR	InterPro: IPR000719; Euk_pkinase.			
DR	InterPro: IPR002290; Ser_thr_kin_actsite.			
DR	Pfam: PF00069; pkinase; 1.			
DR	SMART: SMO0220; S_TKRC; 1.			
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE: PS00101; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE: PS00108; PROTEIN_KINASE_SF; 1.			
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.			
SQL	SEQUENCE 455 AA; 51566 MW; 12046F84A363F7A0 CRC64;			
Query Match	41.7%;	Score 758.5;	DB 4;	Length 455;
Best Local Similarity	44.0%;	Pred. No. 1.7e-58;		
Matches	150;	Conservative	61;	Mismatches 91;
				Indels 39;
				Gaps
Qy	1	MEYKELAGTGGSSGVYFKCNKTSQGVYAKKFKVESDDPVKKIALREIKLQKLK	60	
Db	1	MENYETLKGVEGSGYGVKKKNTQCIQVAKIETPERQSS-VNKIAMREIKFLQFHH	59	
Qy	61	PNLVNLIEVFRKKRKMHLVFEYCDHTLLNELRPNGVAGDVKISVLQWTLQALNFCHTH	120	
Db	60	ENLVNLIEVFRQKKKHLVFEFIDHTVDELQHYCHGLESKRLKYLQILRAIDYLSHN	119	
Qy	121	NCTHRDKPENLITKGIKICDPGFAQIL-IPGDATDYVATRWYRAPELLVGDQYQ	179	
Db	120	NILHRDKPENILVSQSGITKLCDGFGARTLAAPGDIVTDVATRWYRAPELLVKDTSYQ	179	
Qy	180	SSVDIWAICGVPAELTGQLPWPKGSVDQOYLILRTLVETGFHHVDQAGLELITSSDPP	239	
Db	180	KPVDIWAIGCMIIEMATGNPYLPSSSDLLDLHLVLKV-----	217	
Qy	240	AVAS0AGITGKLI	PHQSIKFSNGFHGISIPBEDMETLEKFSVDHPVALNFMKG	299
Db	218	-----GNLSPLQNIFSKSIFAGVVLQVQHPKNAKKYPKINGLLADIVHACL	267	
Qy	300	KMNPPDRILTC	SOLLESSYF--DSFOE---AQIKRKARNEGR	335
Db	268	QIDPDIRISSOLLHHEFTYRDFGTEKFWPELKAQLLEAK	308	
RESULT	9			
Q9BMG2				

```
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 1106 AA; 118884 MW; 24AC26F5C209923E CRC64;

Query Match 40.9%; Score 743.5; DB 5; Length 1106;
Best Local Similarity 43.7%; Pred. No. 1.3e-56;
Matches 145; Conservative 62; Mismatches 88; Indels 37; Gaps 3;

Qy 1 MEKYEKLAKTSGSGYGVVFCRNKTSQGVAVKKEVESEDDPVVKKIALRMLKQLKH 60
Dy 1 MEAYETLGLGEGTYGVVVKARSVTGKLVARFKQTEQDEHVRKTSSEVRMLQLQH 60
Qy 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNELRNPNGVADGVKISVLWQTLQALNFCCHI 120
Dy 61 PNVIRLEDFRRGKLYLVFEFDHTLTLLSTTRGFHRHRLRYTYQLRGIEFCHQ 120
Qy 121 NCIHDRKIPENILITKOGIKICDFGFAQILIPGDAYTDYVATRWYRAPELVGDTQYG 180
Dy 121 NTIHRDKVPENLVIDESGLLKLCDFGARQTSAGKYTDYVATRWYRAPELVGVAYGK 180
Qy 181 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP 240
Dy 181 PVDVWALGCMFAELSDGQPLFGESDLDQLCLIMQT----- 216
Qy 241 VASOSAGITGKLIPRHQSIFKSNGFHGISIPEPEDEMTELEKFSVDVHPVALNFMKGL 300
Dy 217 -----CGVPQRLVFIFMHPNLYNGISFPHTDILYTLKDYHRESNDWIEFLSSCLH 268
Qy 301 MNPDRILTSQLESSEYF--DSFQ---EAQIK 327
Dy 269 TDPAQRILCTELMELPYTRDGRFDRYEALR 300

RESULT 11
Q9JM02
ID Q9JM02 PRELIMINARY; PRT; 457 AA.
AC Q9JM02;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE SERINE/THRONINE KINASE NKIATRE BETA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Haq R., Randall S., Midmer M., Iafate B., Zanke B.W.;
RT "Cloning and functional characterization of NKIATRE, a novel kinase
related to both MAPKs and cyclin-dependent kinases.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF112183; AAF34870.1; .
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 457 AA; 51894 MW; 99D0782794099CDA CRC64;
```

```
Query Match 40.5%; Score 737.5; DB 11; Length 457;
Best Local Similarity 42.5%; Pred. No. 1.2e-56;
Matches 145; Conservative 65; Mismatches 92; Indels 39; Gaps 5;

Qy 1 MEKYEKLAKTSGSGYGVVFCRNKTSQGVAVKKEVESEDDPVVKKIALRMLKQLKH 60
Dy 1 MEAYETLGVGSGYGVVCKKHKTGRIVAIKIFYE-KPEKSVNKIATREIKLKQFRH 59
```

```
Qy 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNELRNPNGVADGVKISVLWQTLQALNFCCHI 120
Dy 60 ENLVNLIEVFRQKKIHLVFEFDHTVLDLQHYCHGLESKRLRYLFOILRAIEYLHNN 119
Qy 121 NCIHDRKIPENILITKOGIKICDFGFAQILIPGDAYTDYVATRWYRAPELVGDTQYG 179
Dy 120 NIIHRDKIPENILVSQSGITKLCDFGARFARTLAAPGDVYTDYVATRWYRAPELVKDTYTG 179
Qy 180 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP 239
Dy 180 KPDVWALGCMFIEMATGNPILPSSSDLDLLHKVLK----- 217
Qy 240 AVASOSAGITGKLIPRHQSIFKSNGFHGISIPEPEDEMTELEKFSVDVHPVALNFMKGL 299
Dy 218 -----GNLTPLHLNIFESKSPFAGVLPQVQHPKNARKKYPKLGILLADIHVHACL 267
Qy 300 KNPDDRLTSQLESSEYF--DSFQ---AQIKRKARNEGR 335
Dy 268 QIDPAERISSTOLLHHDYFTRDGFIEKFIPELRAKLLQEAQ 308

RESULT 12
Q9JM01
ID Q9JM01 PRELIMINARY; PRT; 505 AA.
AC Q9JM01;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE SERINE/THRONINE KINASE NKIATRE ALPHA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Haq R., Randall S., Midmer M., Iafate B., Zanke B.W.;
RT "Cloning and functional characterization of NKIATRE, a novel kinase
related to both MAPKs and cyclin-dependent kinases.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF112184; AAF34871.1; .
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 505 AA; 57242 MW; 89747921FF2CE683 CRC64;
```

```
Query Match 40.5%; Score 737.5; DB 11; Length 505;
Best Local Similarity 42.5%; Pred. No. 1.4e-56;
Matches 145; Conservative 65; Mismatches 92; Indels 39; Gaps 5;
```

```
Qy 1 MEKYEKLAKTSGSGYGVVFCRNKTSQGVAVKKEVESEDDPVVKKIALRMLKQLKH 60
Dy 1 MEAYETLGVGSGYGVVCKKHKTGRIVAIKIFYE-KPEKSVNKIATREIKLKQFRH 59
Qy 61 PNLVNLIEVFRKKRMHLVFEYCDHTLLNELRNPNGVADGVKISVLWQTLQALNFCCHI 120
Dy 60 ENLVNLIEVFRQKKIHLVFEFDHTVLDLQHYCHGLESKRLRYLFOILRAIEYLHNN 119
Qy 121 NCIHDRKIPENILITKOGIKICDFGFAQILIPGDAYTDYVATRWYRAPELVGDTQYG 179
Dy 120 NIIHRDKIPENILVSQSGITKLCDFGARFARTLAAPGDVYTDYVATRWYRAPELVKDTYTG 179
Qy 180 SVDIWAIGCVFAELLTGQPLWPGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPP 239
Dy 180 KPDVWALGCMFIEMATGNPILPSSSDLDLLHKVLK----- 217
Qy 240 AVASOSAGITGKLIPRHQSIFKSNGFHGISIPEPEDEMTELEKFSVDVHPVALNFMKGL 299
```



us-09-671-050-10.rspt

Fri May 3 11:22:26 2002

```
Db 218 -----GNLPHLNFESKPIFAGVVLPOVQHPKNAKKYKPLNGLLADIHVACL 267
Qy 300 KMPDRRLTCQSLESYF--DSFQB---AQIKRKARNEGR 335
Db 268 QIDPAERISDLDLHHDYFTRDGFIEKFTEPELRAKLLOEAK 308

RESULT 13
Q9W6R6 PRELIMINARY; PRT; 1104 AA.
AC Q9W6R6;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE SERINE-THREONINE KINASE 9.
GN STK9.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99263230; PubMed=10330123;
RA Brunner B., Todt T., Lenzner S., Stout K., Schulz U., Ropers H.-H.,
RA Kalscheuer V.M.;
RT "Genomic structure and comparative analysis of nine Fugu genes:
RT conservation of synteny with human chromosome Xp22.2-p22.1.";
RL Genome Res. 9:437-448(1999).
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF146687; AAD28798.1; -.
DR HSP; Q16539; IWF.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW KW
SQ SEQUENCE 1104 AA; 123751 MW; 9085A078B551D876 CRC64;

Query Match 34.0%; Score 618; DB 13; Length 1104;
Best Local Similarity 38.0%; Pred. No. 1.4e-45;
Matches 128; Conservative 66; Mismatches 97; Indels 46; Gaps
7;

Qy 24 KTSQVAVKFESEDDPVVKKIALREIRMLKQKHPNLVNLIEVFRKRKMHLYFEYC 83
Db 19 KETNELVAIKFKDSENEEVKETTLLRELKMLRTLKQDNIVELKEAFRRGKLYLVEYV 78
Qy 84 DHTLNELEERNPVGAVIKSVLMOTLQALNFCHIHNCIHRDIKPENLITKOGIIC 143
Db 79 ERNMLELLEEHPTGAPPDKVRSYIYQLIKAINWCNKNEIVHRDIKPENLITSSD 138
Qy 144 DFGFAQLIPG-DA-YTDYVATRYRAPPELLVGTQYGVSSDIWAIGCVFAELLTGP 201
Db 139 DFGFARNLEGGDANITYEVATRYRSPPELLLG-APYKAVDMWSVGCILGELS 197
Qy 202 PGKSDVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPAVASQASAGITGKLI 261
Db 198 PGSEIDQLFTQKYL-----GFLPSEQMKLFYSNPR 225
Qy 262 SNGPFHGISIPPEDEMTELEEFSDV-HPVALNFMKGCCKMNPDDRLTCSQLES 320
Db 226 NNPFHGIRPSPVTHPTQLERRYGLSLGLMLDLKMLNLLNPTERYLTQSLNHP 285
Qy 321 FOEAQIRKKA-----RNEGRNRR-----QQVILPKS 347
Db 286 LRQVERERAPPASPNNPRSKRTHHGHENTVPTRS 322
```

```
RESULT 14
Q9UJL6 PRELIMINARY; PRT; 997 AA.
AC Q9UJL6;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE DJ245G19.2 (SERINE/THREONINE KINASE 9) (FRAGMENT).
GN STK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; Z92542; CAB38870.1; -.
DR HSP; P28523; IAG0.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000508; Peptidase_S26.
DR InterPro; IPR002290; Ser_thr_kin_actsite.
DR Pfam; PF00069; pkinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00501; SPASE_I.1; UNKNOWN.1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW NON_TER 1
FT SQ SEQUENCE 997 AA; 111939 MW; B922855AED72608D CRC64;

Query Match 33.8%; Score 615.5; DB 4; Length 997;
Best Local Similarity 39.0%; Pred. No. 2.1e-45;
Matches 124; Conservative 65; Mismatches 90; Indels 39; Gaps
5;

Qy 28 QVAVKFESEDDPVVKKIALREIRMLKQKHPNLVNLIEVFRKRKMHLYFEYCDHTL 87
Db 4 EIVAIKFKDSENEEVKETTLLRELKMLRTLKQDNIVELKEAFRRGKLYLVEYK 63
Qy 88 LNELEERNPVGAVIKSVLMOTLQALNFCHIHNCIHRDIKPENLITKOGIICDGF 147
Db 64 LELLEEMPNGVPPEKYSYIYQLIKAINWCNKNDIVHRDIKPENLITSHNDVLK 123
Qy 148 AOLILPGD--AYTDYVATRYRAPPELLVGTQYGVSSDIWAIGCVFAELLTGP 205
Db 124 ARNLSEGNANYTEVATRYRSPPELLLG-APYKSVDMWSVGCILGELSDG 182
Qy 206 DVDQLYLIIRTLVETGFRHVDQAGLELLTSSDPPAVASQASAGITGKLI 265
Db 183 EIDQLFTQKYL-----GFLPSEQMKLFYSNPR 210
Qy 266 FHGISIPPEDEMTELEEFSDV-HPVALNFMKGCCKMNPDDRLTCSQLES 324
Db 211 FHGLRFPVNHPTQLERRYGLSLGLMLDLKMLNLLKLPADRYLTEQC 267
Qy 325 QIKRKARNEGRNRKQOV 342
Db 268 RLDRSPRSKRPPYHV 285

RESULT 15
Q9ZR11 PRELIMINARY; PRT; 294 AA.
AC Q9ZR11;
DT 01-MAY-1999 (TremBLrel. 10, Created)
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE P34CDC2.
GN CDC2TAA.
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
```

OC	Triticeae; Triticum.
OX	NCBI_TaxID=4565;
RN	[1]
RA	SEQUENCE FROM N.A.
RP	Dong C., John P.C.L.;
RR	Submitted (MAR-1995) to the EMBL/GenBank/DDJB databases.
RL	- !- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
RC	EMBL: U23409; A010483.1; -.
DR	HSSP: P24941; LJSD.
DR	Mendel; 37149; Triae;2321;37149.
DR	InterPro: IPR000719; Euk_pkinase.
DR	InterPro: IPR002290; Ser_thr_kin_actsite.
DR	Pfam: PF00629; pkinase; 1.
DR	SMART: SM00260; S_TKC; 1.
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW	ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ	SEQUENCE 294 AA; 33996 MW; 77ABC487FA862A72 CRC64;

[illegible]

Search completed: May 3, 2002, 11:08:12  
Job time: 351 sec

us-09-671-050-10-rai

Fri May 3 11:22:24 2002

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

28 489.5 26.9 282 1 US-08-318-947A-19  
29 489.5 26.9 282 2 US-08-795-303-19  
30 489 26.9 323 3 US-08-874-347-21  
31 489 26.9 323 3 US-09-093-522-21  
32 488.5 26.8 317 1 US-08-463-090B-9  
33 488.5 26.8 317 2 US-08-874-347-18  
34 488.5 26.8 317 3 US-09-093-522-18  
35 486 26.7 324 3 US-08-874-347-20  
36 486 26.7 324 3 US-09-093-522-20  
37 475 26.1 274 2 US-07-857-224B-29  
38 470.5 25.9 298 2 US-08-061-636-3  
39 470.5 25.9 298 2 US-08-874-347-19  
40 470.5 25.9 298 3 US-09-093-522-19  
41 470.5 25.9 298 5 PCT-US94-05268-3  
42 469.5 25.8 326 2 US-08-306-511A-10  
43 469.5 25.8 326 2 US-08-893-274-10  
44 467 25.7 316 1 US-08-403-634-4  
45 467 25.7 316 4 US-08-913-441B-4

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 20, Appl  
Sequence 29, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 18, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-08-874-347-25  
; Sequence 25, Application US/08874347  
; Patent No 5863741  
; GENERAL INFORMATION:  
; APPLICANT: Limper, Andrew H.  
; APPLICANT: Leof, Edward B. F.  
; APPLICANT: Thomas, Charles F.  
; APPLICANT: Gustafson, Michael P.  
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
; TITLE OF INVENTION: CARINII  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C., P.A.  
; STREET: 60 South Sixth Street, Suite 3300  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/874,347  
; FILING DATE: 13-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 07039/055001  
; TELEPHONE: 612-335-5070  
; TELEFAX: 612-288-9696  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 298 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-874-347-25

Query Match  
Best Local Similarity 30.6%; DB 2; Length 298;  
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

SUMMARIES

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match %	Length	ID	Description
1	557	30.6	298	2	US-08-874-347-25
2	557	30.6	298	2	US-08-969-106-2
3	557	30.6	298	3	US-09-093-522-25
4	554	30.4	297	1	US-08-176-620A-16
5	554	30.4	297	2	US-08-461-985-16
6	551.5	30.3	294	2	US-08-874-347-26
7	551.5	30.3	294	3	US-09-093-522-26
8	547.5	30.1	270	2	US-07-857-224B-31
9	545	29.9	297	2	US-08-874-347-24
10	545	29.9	297	3	US-09-093-522-24
11	541	29.7	297	2	US-08-874-347-23
12	541	29.7	297	3	US-09-093-522-23
13	540	29.7	274	1	US-08-318-947A-20
14	540	29.7	274	2	US-08-795-303-20
15	538.5	29.6	300	2	US-08-874-347-10
16	538.5	29.6	300	3	US-09-093-522-10
17	526	28.9	292	1	US-08-154-915-2
18	526	28.9	292	2	US-08-464-517-38
19	526	28.9	292	2	US-08-246-361A-38
20	526	28.9	292	3	US-08-463-772-38
21	526	28.9	292	5	PCT-US93-09945-2
22	526	28.9	297	4	US-08-932-787B-21
23	526	28.9	297	4	US-08-932-012C-21
24	526	28.9	297	4	US-08-888-818C-21
25	498.5	27.4	297	2	US-08-874-347-22
26	498.5	27.4	297	3	US-09-093-522-22
27	493.5	27.1	274	2	US-07-857-224B-30

Query Match 30.68; Score 557; DB 2; Length 298;  
Best Local Similarity 38.44; Pred. No. 1.2e-50;  
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;

```

RESULT      3
US-09-093-522-25
; Sequence 25, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustarson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINI
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-093-522-25

```

US-08-176-620A-16

Query Match 30.6%; Score 557; DB 3; Length 298;  
Best Local Similarity 38.4%; Pred. No. 1.2e-50;  
Matches 118; Conservative 54; Mismatches 89; Indels 46; Gaps 6;  
QY 1 MEKYEKATGSGYGVVFCRNKTSQGVAVKKF-VESEDDPVVKKIALREIRMLKOLKH 60  
DB 1 MENFQVEKIGETGYGVVYKARNKLTGEVVALKIRLDTETEGVPSTAIRLSLLKELNH 60  
QY 61 PNLVNLIEVFRKRKMHLYFEYCDHTLLNELEARNP-NGVADGVKSVLWQTLQALNFC 119  
DB 61 PNIVKLLDVTHTENKLYLVPEELHQDLKKFMDASALTGIPLIPLIKSYLFQLLQGLAFCHS 120  
QY 120 HNCIHRDIPENILITKQGIKICDFGAOIL-IPGDAYTDYVATRWYRAPELLVGDTOY 178  
DB 121 HRVLRDLKPNOLLINTEGAIKLADFGLARAFGVRYTHVWTLWYRAPELLGSKYY 180  
QY 179 GSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTL----- 217  
DB 181 STAVDINSLGICPAENVTRRALFPDSEIDQLFRIFRFLTGPDDEVVWPGVTMPDYKPSF 240  
QY 218 ---VETGFRHV---DOAGLELLTSS---DPPAVASQAGITGKLIPRHOSIFKNGFFH 267  
DB 241 PKWAPQDFSKVVPLEDGCRSLLSQMLHYDPNKRISAKAAL-----AHPEFQ 287  
QY 268 GISIEP 274  
DB 288 DVTKPVP 294

RESULT 4  
US-08-176-620A-16  
; Sequence 16, Application US/08176620A  
; Patent No. 5593904  
; GENERAL INFORMATION:  
; APPLICANT: Boulton, Teri G.  
; APPLICANT: Cobb, Melanie H.  
; APPLICANT: Yancopoulos, George D.  
; APPLICANT: Nye, Steven  
; APPLICANT: Panayotatos, Nikos  
; TITLE OF INVENTION: A Family of Map2 Protein Kinases  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/176,620A  
; FILING DATE: 03-JAN-1994  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 6526-123  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 297 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein

US-08-176-620A-16

Query Match 30.4%; Score 554; DB 1; Length 297;  
Best Local Similarity 40.5%; Pred. No. 2.4e-50;  
Matches 118; Conservative 58; Mismatches 73; Indels 42; Gaps 7;  
QY 1 MEKYEKATGSGYGVVFCRNKTSQGVAVKKF-VESEDDPVVKKIALREIRMLKOLKH 59  
DB 1 MEDYTKIEKIGETGYGVVYKGRHKTGQGVAMKIRLSEEEG-VFSTAIRLSLLKELR 59  
QY 60 HPNLVNLIEVFRKRKMHLYFEYCDHTLLNELEARNP--VADGVTKSVLWQTLQALNFC 117  
DB 60 HPNIVSLQDYLQMSDRLYLFEFLSMDLKKYLDSSIPQGYMDSSLVSKSYLQILQGVFC 119  
QY 118 HNCIHRDIPENILITKQGIKICDFGAOIL-IPGDAYTDYVATRWYRAPELLVGD 176  
DB 120 HSRVLRDLKPNOLLINTEGAIKLADFGLARAFGIPRVYTHVWTLWYRSPVLLGSA 179  
QY 177 QGSSVDIWAIGCVFAELLTGQPLWPGKSDVDQYLIIRTL----- 217  
DB 180 RYSTPVDINSGITFAELATKPLFHGDSEIDQLFRIFRFLTGTPNNVWPEVESLQYKN 239  
QY 218 -----VETGFRHVQAGLELLTSS---DPPAVASQAGITGKLIPRH 256  
DB 240 TEPKWKPGSLASHVKNLDEGLDLSLKKLIYDP-----AKRISGKMLNH 284

RESULT 5  
US-08-461-985-16  
; Sequence 16, Application US/08461985  
; Patent No. 5872006  
; GENERAL INFORMATION:  
; APPLICANT: Boulton, Teri G.  
; APPLICANT: Cobb, Melanie H.  
; APPLICANT: Yancopoulos, George D.  
; APPLICANT: Nye, Steven  
; APPLICANT: Panayotatos, Nikos  
; TITLE OF INVENTION: A Family of Map2 Protein Kinases  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,985  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/176,620  
; FILING DATE: 03-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 6526-123  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 297 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown

MOLECULE TYPE: protein  
US-08-461-985-16

```

Query Match      30.4%; Score 554; DB 2; Length 297;
Best Local Similarity 40.5%; Pred. No. 2.4e-50;
Matches 118; Conservative 58; Mismatches 73; Indels 42; Gaps 7;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGOVAVKKF-VESDDDPVKKKTALREIMLKQKH 59
DB 1 MEDYTKIERIGEGTYGVVYKGRHKTGGVAMKKIRLESEEG-VPSTAIRKLSLLKEIR 59

QY 60 HPLVNLIEVFRKRMHLVFECYDHTLLNELRPNNG--VADGVKISVLWQTLQALNFC 117
DB 60 HPNIVSLQDVLWQSRUYLIFELFSLMDLKKYLDISPPGYMDSSLVKSILYQTLQGVFC 119

QY 118 HTHNCIHRDKPENILITKOGIKICDFGFAQIL-IPGDAYDYVATRWYRAPELLVGT 176
DB 120 HSRVLHRDLKPNQLTDDKGTIKLADFLARAFGIPIRVYHYVYVTLWTRSEVLLGSA 179

QY 177 QYGSVDYWAIGCVFAELLTGQPLWPGKSDVDQIYLIRTL----- 217
DB 180 RYSTPVDWSIGTIFAELATKPKLPHGDSIDOLFRIFRALGTPNNEVWPEVESLQDYKN 239

QY 218 -----VETGRHVDQAGLELITSS---DPPAVASQASGITKGLIPRH 256
DB 240 TPKWKPGSLASHVKNLDENGLDLSKMLIYDP-----AKRISGKNALNH 284

```

## RESULT 6

```

US-08-874-347-26
; Sequence 26, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leaf, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-874-347-26

```

```

Query Match      30.3%; Score 551.5; DB 2; Length 294;
Best Local Similarity 34.8%; Pred. No. 4.4e-50;
Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;

QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGOVAVKKFVESEDDPVVKKKTALREIMLKQKH 60
DB 1 MEQYEKEKIGEGTYGVVYKGRHKTGGVAMKKIRLESEEGVPSTAIRKLSLLKEMH 60

QY 61 PNLVNLIEVFRKRMHLVFECYDHTLLN-----ELERNPNGVADGVKISVLWQTLQAL 114
DB 61 GNIVRLHDVILHSEKRYLVFEYLDLIDLKKFMDSCPEFAKNPT-----LIKSYLQILRGV 115

QY 115 NFCHINCHIRDKPENILITKOGIKICDFGFAQIL-IPGDAYDYVATRWYRAPELL 172
DB 116 AYCHSHRVLHRDLKPNQLTDDKGTIKLADFLARAFGIPVTFTHHEVYVTLWYRAPEIL 175

QY 173 VGDYQYSSVDYWAIGCVFAELLTGQPLWPGKSDVDQIYLIRTLVETGTFHRVDAQGLEL 232
DB 176 LGSQYXSTPVDMSVGCIFAEVMVQKPLFGDSEIDELFKIFRVLGTP----- 223

QY 233 LTSSDPPAVASQASGITKGLIPRHQSIFKSNFGFFHGISIPEPEDEMETLEEKSDVHPVAL 292
DB 224 -NEQSWEGVSS-----LPDYKSAFPK-----WQADLATI---VPTLDPAGL 261

QY 293 NFMKGCLKMPDRLTLCSLLESSYDFDSFQEAQ 325
DB 262 DLLSKMLRYEPNKRITARQALEHEYFKDLEMVO 294

RESULT 7
US-09-093-522-26
; Sequence 26, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leaf, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids

```

```

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-093-522-26

Query Match 30.3%; Score 551.5; DB 3; Length 294;
Best Local Similarity 34.8%; Pred. No. 4,4e-50;
Matches 116; Conservative 68; Mismatches 102; Indels 47; Gaps 8;

QY 1 MEYKEKLAKTGEGSYGVFKCRNKTSGOVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 MEQYEKEKIGEGYGVYRDKVTNETIALKIRLROEDEGVPSTAIRISLLKELRH 60

QY 61 PNLVNLIEVFRKKMHLVFECYDHTLLN-----ELENFNGVADGVKSVIMOTLOAL 114
Db 61 GNIVRLHDVTHSEKRIYLVFELDLDLKKFMDSCPEFAKNPT-----LIKSYLYQILRGV 115

QY 115 NPTHCHIRHDKPENILITKO-GIIKICDFGAQIL-IPGDATDYVATRYRAPELL 172
Db 116 AYCHSHRYLHRDLKPNQNLIDRTNALKADGLARAFGIPRVTFHEVVTLYRAPELL 175

QY 173 VGDYQYSSVDIWAIGCVFAELLTQGLPMPGKSDVDOLYLIIRTLVETGFRHVDQAGLE 232
Db 176 LGSROYSTPDMWSVGCIFAEMVQKPLFGDSEIDELFKIFRVLGTP----- 223

QY 233 LTSDDPPAVASOSAGITGKLIPRHOSIFKSNFGPHGISIPEDEMTELEKFSVDHPVAL 292
Db 224 -NEQSMFGVSS-----LPDYKSAPFK-----WQADLATI---VPTLDPAGL 261

QY 293 NFMKGGKMNPDRLTCSQLLESSYFDSFQEAQ 325
Db 262 DLLSKMLRYEPNKRITARQALEHEYFKOLEMVQ 294

RESULT 8
US-087-224B-31
; Sequence 31, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; ORIGINAL SOURCE:

;
; ORGANISM: human
; FEATURE: Protein kinase; Table 8 Column 35
; PUBLICATION INFORMATION:
; AUTHORS: Hanks, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988
;
US-07-857-224B-31

Query Match 30.1%; Score 547.5; DB 2; Length 270;
Best Local Similarity 43.4%; Pred. No. 1e-49;
Matches 118; Conservative 56; Mismatches 73; Indels 25; Gaps 8;

QY 2 ERYEKLAKTGEGSYGVFKCRNKTSGOVAVKKFVESEDDPVVKKIALREIRMLKQLKH 60
Db 1 EDYTKIEKIGEGYGVYGRHKTTGGVAMKKIRLESEEG-VPSTAIRISLLKELRH 59

QY 61 PNLVNLIEVFRKKMHLVFECYDHTLLNELRNPN--VADGVKSVLQTLQALNECH 118
Db 60 PNIVSLQDVLQMODSRLYLIFELFSLMDLKKYLDIPPGQYMDSSLVKSYLQILQGVIFCH 119

QY 119 ITHCHIRHDKPENILITKO-GIIKICDFGAQIL-IPGDATDYVATRYRAPELLVGDQ 177
Db 120 SRVJLHRLKPNQNLIDDDGTIKLADFLARAFGIPRVTFHEVVTLYRSPEVLLGSAR 179

QY 178 YGSSVDIWAIGCVFAELLTQGLPMPGKSDVDOLYLIIRTL-----VETGFRHVDQ 227
Db 180 YSTPVDIWSIGTIFAEELATKKPLFGDSEIDQLFRIFRALGNTNPNVWPEVES-LQOLDE 238

QY 228 AGLEILTSS---DPPAVASOSAGITGKLIPRH 256
Db 239 NGDLLSKMLIYD-----AKRISCKMALNH 264

RESULT 9
US-08-874-347-24
; Sequence 24, Application US/08874347
; Patent No. 5863741
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-874-347-24

Query Match 29.9%; Score 545; DB 2; Length 297;  
Best Local Similarity 40.2%; Pred. No. 2.2e-49;  
Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;  
QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGOVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59  
DB 1 MEDYTKIEKIGEGTYGVVYKGRHKTGGVAMKKIRLESEEG-VPTAIRREISLLKELR 59  
QY 60 HPNLVNLIETVFRKRKMHVFEYCDHTLLNELERNPG--VADGVKISVLWQTLQALNFC 117  
DB 60 HPNIVSLQDVLMDQSLYLIFELSMCLKYLDIPPGQYMDSSLVSKSYLYQILOGIVFC 119  
QY 118 HIHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDT 176  
DB 120 HSRRVLRDLKPNLLIDDKGTIKLADFGARAFGIPIRVYTHEVVTLMWRSPEVLLGSA 179  
QY 177 QYSSVDIWAIGCVFAELITGQPLWPKGSDVDQLYLIIRTL----- 217  
DB 180 RYTPVDIWSIGTIFAEELATKKPLFHGDSIDQLFRIFALGTIPNNEVPEVESLQDYKN 239  
QY 218 -----VETGFRHVDQAGLELITSS---DPPAVASQASAGITGKLIPRH 256  
DB 240 TFPKWKPGSLASHVKNLDENGLDLSKMLIYDP-----AKRISGKMLNH 284

RESULT 10  
US-09-093-522-24  
Sequence 24, Application US/09093522  
Patent No. 6015700  
GENERAL INFORMATION:  
APPLICANT: Limper, Andrew H.  
APPLICANT: Leof, Edward B.  
APPLICANT: Thomas, Charles F.  
APPLICANT: Gustafson, Michael P.  
TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C., P.A.  
STREET: 60 South Sixth Street, Suite 3300  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/093,522  
FILING DATE: 08-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/874,347  
FILING DATE: 13-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055002

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-093-522-24

Query Match 29.9%; Score 545; DB 3; Length 297;  
Best Local Similarity 40.2%; Pred. No. 2.2e-49;  
Matches 117; Conservative 58; Mismatches 74; Indels 42; Gaps 7;  
QY 1 MEKYEKLAKTGGSGYGVVFKCRNKTSGOVAVKKF-VESEDDPVVKKIALREIRMLKQLK 59  
DB 1 MEDYTKIEKIGEGTYGVVYKGRHKTGGVAMKKIRLESEEG-VPTAIRREISLLKELR 59  
QY 60 HPNLVNLIETVFRKRKMHVFEYCDHTLLNELERNPG--VADGVKISVLWQTLQALNFC 117  
DB 60 HPNIVSLQDVLMDQSLYLIFELSMCLKYLDIPPGQYMDSSLVSKSYLYQILOGIVFC 119  
QY 118 HIHNCIHRDIKPENILITKQGIKICDFGFAQIL-IPGDAYTDYVATRWYRAPELLVGDT 176  
DB 120 HSRRVLRDLKPNLLIDDKGTIKLADFGARAFGIPIRVYTHEVVTLMWRSPEVLLGSA 179  
QY 177 QYSSVDIWAIGCVFAELITGQPLWPKGSDVDQLYLIIRTL----- 217  
DB 180 RYTPVDIWSIGTIFAEELATKKPLFHGDSIDQLFRIFALGTIPNNEVPEVESLQDYKN 239  
QY 218 -----VETGFRHVDQAGLELITSS---DPPAVASQASAGITGKLIPRH 256  
DB 240 TFPKWKPGSLASHVKNLDENGLDLSKMLIYDP-----AKRISGKMLNH 284

RESULT 11  
US-08-874-347-23  
Sequence 23, Application US/08874347  
Patent No. 5863741  
GENERAL INFORMATION:  
APPLICANT: Limper, Andrew H.  
APPLICANT: Leof, Edward B.  
APPLICANT: Thomas, Charles F.  
APPLICANT: Gustafson, Michael P.  
TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C., P.A.  
STREET: 60 South Sixth Street, Suite 3300  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/874,347  
FILING DATE: 13-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-335-5070



us-09-671-050-10.ra1

Fri May 3 11:22:24 2002

TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-093-522-23

Query Match 29.7%; Score 541; DB 2; Length 297;  
Best Local Similarity 39.6%; Pred. No. 5.8e-49;  
Matches 114; Conservative 61; Mismatches 77; Indels 36; Gaps 6;

QY 1 MEKYEKLAKTGGSGVVFCKRNKTSQVVAVKPF-VESEDDPVVKKIALRMLKOLK 59  
DB 1 MEDIKIEKIGEGTGVVYKGRHTTGOIVAMKKIRLESEEG-VPSTAIRISLLKELR 59  
QY 60 HPNVLNIEVFRKRRKMHLEFYCDHTLLNELRNPNP--VADGVKSVLWOTLQALNFC 117  
DB 60 HPNIVSLQDYLMDQSLXLIIEFLSMDLKKYLDSPGQFMDSSLKSVLYQILOGIVFC 119  
QY 118 HTHNCIHRDKPENILITKQIIKICDFGFAQIL-IPGDAYTDVATRWYRAPELLVGD 176  
DB 120 HSRRLVHRDLKPNQLLIDDKGTIKLADFGLARAFGIPRVVTHEVVTWYRSPEVLLGSA 179  
QY 177 QYGSVDIWAIGCVFAELLTGQPLWPKSDVDOLYLITL----- 217  
DB 180 RYSTPVDIWSIGTIFAELATKPLPHGDSIDQLFRIFALGTPTNNEVWEVESLQDYKN 239  
QY 218 -----VETGFRHVDQAGLELITSDPPAVASQASAGITGKLIPIRH 256  
DB 240 TFPKWKPGSLASHVKNLDENGLDLSKM--LVYDPAKRISGKMAKH 284

RESULT 12  
US-09-093-522-23  
Sequence 23, Application US/09093522  
Patent No. 6015700  
GENERAL INFORMATION:  
APPLICANT: Limper, Andrew H.  
APPLICANT: Leaf, Edward B.  
APPLICANT: Thomas, Charles F.  
APPLICANT: Gustafson, Michael P.  
TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
TITLE OF INVENTION: CARINII  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C., P.A.  
STREET: 60 South Sixth Street, Suite 3300  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/093,522  
FILING DATE: 08-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/874,347  
FILING DATE: 13-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055002  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-093-522-23

Query Match 29.7%; Score 541; DB 3; Length 297;  
Best Local Similarity 39.6%; Pred. No. 5.8e-49;  
Matches 114; Conservative 61; Mismatches 77; Indels 36; Gaps 6;

QY 1 MEKYEKLAKTGGSGVVFCKRNKTSQVVAVKPF-VESEDDPVVKKIALRMLKOLK 59  
DB 1 MEDIKIEKIGEGTGVVYKGRHTTGOIVAMKKIRLESEEG-VPSTAIRISLLKELR 59  
QY 60 HPNVLNIEVFRKRRKMHLEFYCDHTLLNELRNPNP--VADGVKSVLWOTLQALNFC 117  
DB 60 HPNIVSLQDYLMDQSLXLIIEFLSMDLKKYLDSPGQFMDSSLKSVLYQILOGIVFC 119  
QY 118 HTHNCIHRDKPENILITKQIIKICDFGFAQIL-IPGDAYTDVATRWYRAPELLVGD 176  
DB 120 HSRRLVHRDLKPNQLLIDDKGTIKLADFGLARAFGIPRVVTHEVVTWYRSPEVLLGSA 179  
QY 177 QYGSVDIWAIGCVFAELLTGQPLWPKSDVDOLYLITL----- 217  
DB 180 RYSTPVDIWSIGTIFAELATKPLPHGDSIDQLFRIFALGTPTNNEVWEVESLQDYKN 239  
QY 218 -----VETGFRHVDQAGLELITSDPPAVASQASAGITGKLIPIRH 256  
DB 240 TFPKWKPGSLASHVKNLDENGLDLSKM--LVYDPAKRISGKMAKH 284

RESULT 13  
US-08-318-947A-20  
Sequence 20, Application US/08318947A  
Patent No. 5798245  
GENERAL INFORMATION:  
APPLICANT: Anderson, Paul J.  
APPLICANT: Tian, Qingsheng  
TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED  
TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sughruue, Mion, Zinn, Macpeak & Seas  
STREET: 2100 Pennsylvania Avenue, NW Suite 800  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,947A  
FILING DATE: 06-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/133,530  
FILING DATE: 07-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mack, Susan J.  
REGISTRATION NUMBER: 30,951  
REFERENCE/DOCKET NUMBER: A6462  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060

TELEFAX: (202)293-2920  
TELEX: 6491103  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 274 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-318-947A-20

Query Match 29.7%; Score 540; DB 1; Length 274;  
Best Local Similarity 46.6%; Pred. No. 6.5e-49;  
Matches 102; Conservative 47; Mismatches 68; Indels 2; Gaps 2;

QY 1 MEKYEKLAKTEGSGYGVFKCRNKTSGOVAVKFFVESEDDPWKKIALREIRMLKOLKH 60  
DB 1 MENFOKVEKIGEGTYGVYKARNKLTGEVVALKKIRLDTETEGVPTAIRISLLKELNH 60  
QY 61 PNLVNLIEVFRKKRKHLYFEYCDHTLLNELRNP-NGVADGVKSVIMOTLOALNFCHI 119  
DB 61 PNLVNLIEVFRKKRKHLYFEYCDHTLLNELRNP-NGVADGVKSVIMOTLOALNFCHI 119  
QY 120 HNCIHRDKPENILITKQIIKICDFGFAQL-IPGDAYTDYVATRWYRAPELLVGDTOY 178  
DB 121 HRVLRDLKPNLLINTEGAIKLADFGARAFGVPVRYTTHVVTWYRAPELLVGSYY 180  
QY 179 GSSVDINAIGCVFAELLTGQPLWPKGSDVDOLYLIIRTL 217  
DB 181 STAVKINSGLCIFAENVTRRALFPDSEIDQLFRIFRTL 219

## RESULT 14

US-08-795-303-20  
Sequence 20, Application US/08795303  
Patent No. 5948656  
GENERAL INFORMATION:  
APPLICANT: Anderson, Paul J.  
APPLICANT: Tian, Qingsheng  
TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED  
TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sudhru, Mion, Zinn, Macpeak & Seas  
STREET: 2100 Pennsylvania Avenue, NW Suite 800  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,303  
FILING DATE: 04-FEB-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/318,947  
FILING DATE: 06-OCT-1994  
APPLICATION NUMBER: 08/133,530  
FILING DATE: 07-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mack, Susan J.  
REGISTRATION NUMBER: 30,951  
REFERENCE/DOCKET NUMBER: A6462  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-2920  
TELEX: 6491103  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 274 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-318-947A-20

LENGTH: 274 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-795-303-20

Query Match 29.7%; Score 540; DB 2; Length 274;  
Best Local Similarity 46.6%; Pred. No. 6.5e-49;  
Matches 102; Conservative 47; Mismatches 68; Indels 2; Gaps 2;

QY 1 MEKYEKLAKTEGSGYGVFKCRNKTSGOVAVKFFVESEDDPWKKIALREIRMLKOLKH 60  
DB 1 MENFOKVEKIGEGTYGVYKARNKLTGEVVALKKIRLDTETEGVPTAIRISLLKELNH 60  
QY 61 PNLVNLIEVFRKKRKHLYFEYCDHTLLNELRNP-NGVADGVKSVIMOTLOALNFCHI 119  
DB 61 PNLVNLIEVFRKKRKHLYFEYCDHTLLNELRNP-NGVADGVKSVIMOTLOALNFCHI 119  
QY 120 HNCIHRDKPENILITKQIIKICDFGFAQL-IPGDAYTDYVATRWYRAPELLVGDTOY 178  
DB 121 HRVLRDLKPNLLINTEGAIKLADFGARAFGVPVRYTTHVVTWYRAPELLVGSYY 180  
QY 179 GSSVDINAIGCVFAELLTGQPLWPKGSDVDOLYLIIRTL 217  
DB 181 STAVKINSGLCIFAENVTRRALFPDSEIDQLFRIFRTL 219

## RESULT 15

US-08-874-347-10  
Sequence 10, Application US/08874347  
Patent No. 5863741  
GENERAL INFORMATION:  
APPLICANT: Limper, Andrew H.  
APPLICANT: Leof, Edward B.  
APPLICANT: Thomas, Charles P.  
APPLICANT: Gustafson, Michael P.  
TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
TITLE OF INVENTION: CARINII  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C., P.A.  
STREET: 60 South Sixth Street, Suite 3300  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/874,347  
FILING DATE: 13-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 300 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-874-347-10

Fri May 3 11:22:24 2002

Query Match 29.6%; Score 538.5; DB 2; Length 300;  
Best Local Similarity 34.6%; Pred. No. 1.1e-48;  
Matches 113; Conservative 76; Mismatches 95; Indels 43; Gaps 9;

QY	1	MERYELAKTGGSGVGVKCRNKTSQGVAVVAVKKFEVESEDDPVVKKIALREIRMLKOLKH	60
DB	1	MEQVORLEKIGEGTGVVYKAKDLESGTIVALKTRLEAEDEGVPSTAIRISLLKEMHN	60
QY	61	PNLVNLEIEFRKRKMHVLVEYCDHTLLNLELERNPVGADG--VKSVLWOTLOALNECH	118
DB	61	DNVYRLNIIHOESRLYLVEFLDLDLKKYMSIPKDMMLGAEMIKKFWSQLVSGVKYCH	120
QY	119	IHNCIHRDKPENILITKOGIIKICDFGAQIL-IPGDAYTDYVATRWYRAPELLVGTQ	177
DB	121	SHRILHRDLKPQNLLIDREGNKLADFGLARAFGVPLRGYTHEVVTLMYRAPEVLLGGRQ	180
QY	178	YGSSVDIWAIGCVFAELLTGQPLWPKSDVDOLYLIIRTLVETGRHVDAQLELLTSSD	237
DB	181	YATALDIWSIGCIFAEMATKKKLPFGDSEIDEIFRFR-----ILGTPD	224
QY	238	PPAVASQAGITGKLIPRHQSIFKSNGFHGISIPE--PEDM-ETLEEKFSVHPVALNF	294
DB	225	----ENSWPGITS--YPDFKATF-----PKWSFKNLGELITELDSD-----GIDL	263
QY	295	MKGCLKMNPDDRLLTCSQLLESSYFDSF	321
DB	264	LQKCLRYYPARISAKKALDHPYFDDF	290

Search completed: May 3, 2002, 11:02:44  
Job time: 78 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 16:29:21 ; Search time 2919.35 Seconds  
(without alignments)  
5340.173 Million cell updates/sec

Title: US-09-671-050-11  
Perfect score: 945  
Sequence: 1 atgggaagatgataaaatt.....aggtaactccgctcaaaagt 945

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- GenBank:\*
- 1: gb\_ba:\*
  - 2: gb\_hgt:\*
  - 3: gb\_in:\*
  - 4: gb\_om:\*
  - 5: gb\_ov:\*
  - 6: gb\_pat:\*
  - 7: gb\_ph:\*
  - 8: gb\_pi:\*
  - 9: gb\_pi:\*
  - 10: gb\_ro:\*
  - 11: gb\_sts:\*
  - 12: gb\_sy:\*
  - 13: gb\_un:\*
  - 14: gb\_vi:\*
  - 15: em\_ba:\*
  - 16: em\_fun:\*
  - 17: em\_hum:\*
  - 18: em\_in:\*
  - 19: em\_om:\*
  - 20: em\_or:\*
  - 21: em\_ov:\*
  - 22: em\_pat:\*
  - 23: em\_ph:\*
  - 24: em\_pl:\*
  - 25: em\_ro:\*
  - 26: em\_sts:\*
  - 27: em\_sy:\*
  - 28: em\_un:\*
  - 29: em\_vi:\*
  - 30: em\_htgo\_hum:\*
  - 31: em\_htgo\_inv:\*
  - 32: em\_htgo\_rod:\*
  - 33: em\_htg\_hum:\*
  - 34: em\_htg\_inv:\*
  - 35: em\_htg\_rod:\*
  - 36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	945	100.0	945	6	AX107722	AX107722 Sequence
2	929	98.3	972	6	AX107716	AX107716 Sequence
3	863	91.3	1083	6	AX166534	AX166534 Sequence
4	839	88.8	1041	6	AX107720	AX107720 Sequence
5	823	87.1	1068	6	AX107714	AX107714 Sequence
6	643	68.0	911	6	AX056404	AX056404 Sequence
7	559.2	59.2	2615	6	AX056405	AX056405 Sequence
8	507	53.7	1819	6	AX107724	AX107724 Sequence
9	454.4	48.1	561	6	AX107712	AX107712 Sequence
10	454.4	48.1	594	6	AX107718	AX107718 Sequence
11	413.8	43.8	1363	9	HSSHPKAB	X66358 H sapiens m
12	287	30.0	3153	4	AB029045	AB029045 Oryctolag
13	283.8	30.0	1993	9	HSU35146	U35146 Human p56 K
14	272.6	28.8	1518	10	AB029067	AB029067 Mus muscu
15	272.6	28.8	3988	10	AB029066	AB029066 Mus muscu
16	272.6	28.8	4022	10	AB029065	AB029065 Mus muscu
17	227.4	24.1	1513	9	AF130372	AF130372 Homo sapi
18	227.4	24.1	1773	6	AX166533	AX166533 Sequence
19	218.4	23.1	101340	2	AC018104	AC018104 Drosophil
20	218.4	23.1	190668	3	AC007419	AC007419 Drosophil
21	218.4	23.1	259973	3	AE003611	AE003611 Drosophil
22	211.4	22.4	1732	10	AF112183	AF112183 Rattus no
23	211.4	22.4	1738	10	AF112184	AF112184 Rattus no
24	208.4	22.1	1974	3	AF326965	AF326965 Trypanoso
25	200.4	21.2	80362	9	AC079615	AC079615 Homo sapi
26	200.4	21.2	155028	33	AC037489	AC037489 Homo sapi
27	200.4	21.2	164281	2	AC092672	AC092672 Homo sapi
28	194	20.5	3399	9	HSTXF3	V15057 Homo sapien
29	180.2	19.1	3321	3	LME293284	AJ293284 Leishmani
30	175.8	18.6	1866	6	AX056407	AX056407 Sequence
31	171.8	18.2	903	8	AF026545	U69958 Trypanosoma
32	168	17.8	960	3	TCU69958	L00652 Dictyosteli
33	163.6	17.3	921	3	DDICRP	AR087351 Sequence
34	163.4	17.3	1476	6	HS087351	X61622 H.sapiens C
35	163.4	17.3	1476	6	HSCDK2MR	AB4497 Sequence 11
36	163.4	17.3	1635	6	A84497	AB4499 Sequence 11
37	163.4	17.3	1635	6	A84499	X62071 H.sapiens C
38	163	17.2	1297	9	HSCDK2	M68520 Human cdc2-
39	161.8	17.1	2213	9	HUMCDC2A	BC003065 Homo sapi
40	161.8	17.1	2264	9	BC003065	AF076534 Cryptospo
41	157.2	16.6	885	3	AF076534	AJ223949 Crictetulu
42	157	16.6	2213	10	CGR223949	X05360 Human CDC2
43	155.6	16.5	1050	9	HSCDC2	Y00272 Human cell
44	155.6	16.5	1050	9	HSCDC2R	U63337 Mus musculu
45	153.8	16.3	1708	10	MMU63337	

ALIGNMENTS

RESULT	1
AX107722	
LOCUS	AX107722 945 bp DNA PAT 30-APR-2001
DEFINITION	Sequence 11 from Patent WO0123579.
ACCESSION	AX107722
VERSION	AX107722.1 GI:13923202
KEYWORDS	human.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 945)
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE	Human kinase proteins and polynucleotides encoding the same
JOURNAL	Patent: WO 0123579-A 11 05-APR-2001;
FEATURES	Lexicon Genetics Incorporated (US)
	Location/Qualifiers
	1..945
	/organism="Homo sapiens"

BASE COUNT		323 a		165 c		202 g		255 t	
ORIGIN									
Query Match		100.0%; Score 945; DB 6; Length 945;							
Best Local Similarity		100.0%; Pred No. 2.2e-202;							
Matches 945; Conservative		0; Mismatches 0; Indels 0; Gaps 0;							
QY	1	atggaagaatgataaaaattagctaagaactggagaagggcttattggggtgtattcaaa	60						
Db	1	ATGGAAGAAGTATGAATAATTAGCTTAAGACTGGAGAAGGCTTATGGGTTGTATTCAAA	60						
QY	61	tgcagaacaaacacctctggacaagtagctgttaaaaaatttgtggaatctgaagat	120						
Db	61	TGCAGAAACAAACCTCTGGACAAGTAGTGTGTTAAAAAATTCTGGAATCTGAAGAT	120						
QY	121	gatcctgttgttaagaaaatagcactaagagaatacgtatgttgaagcaattaaaaacat	180						
Db	121	GATCCTGTGTGTTAAGAAANTAGCAGCTAAGAGAAATACGTATGTGAAAGCAATTAAAAACAT	180						
QY	181	ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt	240						
Db	181	CCAAATCTGTGTGAACCTCATCGAGGTGTTCAGAGAGAAAAAGGAAATGCAATTAGTTT	240						
QY	241	gaatactgtgatacacaacttttaaatgagctgaaagaacccaaaatggagttgctgat	300						
Db	241	GAATACTGTGATCATACACTTTTAATGAGCTGGAAAGAAACCCAAATGGAGTGTCTGAT	300						
QY	301	ggagtgataaaaacgctattatggcaaaccttcaagctcttcaattttctgcatatacat	360						
Db	301	GGAGTGATCAAAAGCGTATTATGGCAAAACCTTCAAGCTCTTAATTTCTGTGATATACAT	360						
QY	361	aactgtattcacagagataaaaacctgaaataatttcaataactaagcaaggaataatc	420						
Db	361	AACTGTATTTCACAGAGATATAAACCTGAAATATTTCTAATAACTAAGCAAGGAATAATC	420						
QY	421	aagattgtcgactcgggtttgcacaaattctgtccaggagatgcctacacggattat	480						
Db	421	AAAGTTTGTGACTTCGGGTTTCACAAATTTCTGATTCTCCAGGAGATGCCTACACCGATTAT	480						
QY	481	gtagtcagagatggtaccgagctcctgaactctctgtggagagatacagataggtct	540						
Db	481	GTAGTCAGAGATGGTACCGAGCTCTTAAATGAGCTGGAAAGAACCCAAATGGAGTTGTTCT	540						
QY	541	tcagtcgataataggggtatggtgtgtttttgtcagagctcctgacaggccagccactg	600						
Db	541	TCAGTCGATATATGGGCTATTTGGTTGTGTTTTTCAGAGCTCCTGACAGCCAGCCACTG	600						
QY	601	atcccgagacataaatcaacttttaaaagttaacgggtttttccatggcagataacct	660						
Db	601	TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAAATTA	660						
QY	661	atcccgagacataaatcaacttttaaaagttaacgggtttttccatggcagataacct	720						
Db	661	ATCCCAAGACATCAATCAATCTTTAAAGTATACGGGTTTTTCCATGGCATCAGTATACCT	720						
QY	721	gagccagaacacatgaaactcttgaggaaaagtctcagatgttcatcctgtggtctgt	780						
Db	721	GAGCCAGACACATGGAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGCTCTG	780						
QY	781	aacttcatgaaggggtgtctgaagatgaatccagatgacagataaacctgtgtcccaactc	840						
Db	781	AACTTCTGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTACCTGTGTCCCAACTC	840						
QY	841	ctggagagctcctacttcttatttttaagagcccaattaaaaagaagaagcagtaat	900						
Db	841	CTGGAGAGCTCCTACTTTGATTCTTTTCAAGAGGCCCAATTTAAAGAAAAAGCACGTAAT	900						
QY	901	gaaggagaacacagagacgcacaggttaacttcgcgtcaaaagt	945						
Db	901	GAAGGAGAAAAACAGAACGCCAACAGGTACTTCCGCTCAAAAGT	945						

RESULT 2	AX107716	972 bp	DNA	PAT	30-APR-2001
LOCUS	AX107716	Sequence 5 from Patent WO0123579.			
DEFINITION	AX107716				
ACCESSION	AX107716				
VERSION	AX107716.1	GI:13923199			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.				
TITLE	Human kinase proteins and polynucleotides encoding the same				
JOURNAL	Patent: WO 0123579-A 5 05-APR-2001;				
FEATURES	Lexicon Genetics Incorporated (US)				
source	Location/Qualifiers				
	1..972				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	323 a	172 c	214 g	263 t	
ORIGIN					
Query Match					
Best Local Similarity 100.0%; Pred. No. 8.7e-199;					
Matches 929; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	atggaagaatgataaaaattagctaagaactggagaagggcttattggggtgtattcaaa	60		
Db	1	ATGGAAGAAGTATGAATAATTAGCTTAAGACTGGAGAAGGCTTATGGGTTGTATTCAAA	60		
QY	61	tgcagaacaaacacctctggacaagtagctgttaaaaaatttgtggaatctgaagat	120		
Db	61	TGCAGAAACAAACCTCTGGACAAGTAGTGTGTTAAAAAATTGTGGAATCTGAAGAT	120		
QY	121	gatcctgttgttaagaaaatagcactaagagaaaacacgtatgttgaagcaattaaaaacat	180		
Db	121	GATCCTGTGTGTTAAGAAANTAGCAGCTTAAGAGAAATACGTATGTGAAAGCAATTAAAAACAT	180		
QY	181	ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt	240		
Db	181	CCAAATCTGTGTAACCTCATCGAGGTGTTCAGAGAGAAAAAGGAAATGCAATTAGTTT	240		
QY	241	gaatactgtgatacacaacttttaaatgagctgaaagaacccaaaatggagttgctgat	300		
Db	241	GAATACTGTGATCATACACTTTTAATGAGCTGGAAAGAAACCCAAATGGAGTGTCTGAT	300		
QY	301	ggagtgataaaaacgctattatggcaaaccttcaagctcttcaattttctgcatatacat	360		
Db	301	GGAGTGATCAAAAGCGTATTATGGCAAAACCTTCAAGCTCTTAATTTCTGTGATATACAT	360		
QY	361	aactgtattcacagagataaaaacctgaaataatttcaataactaagcaaggaataatc	420		
Db	361	AACTGTATTTCACAGAGATATAAACCTGAAATATTTCTAATAACTAAGCAAGGAATAATC	420		
QY	421	aagattgtcgactcgggtttgcacaaattctgtccaggagatgcctacacggattat	480		
Db	421	AAAGTTTGTGACTTCGGGTTTCACAAATTTCTGATTCTCCAGGAGATGCCTACACCGATTAT	480		
QY	481	gtagtcagagatggtaccgagctcctgaactctctgtggagagatacagataggtct	540		
Db	481	GTAGTCAGAGATGGTACCGAGCTCTTAAATGAGCTGGAAAGAACCCAAATGGAGTTGTTCT	540		
QY	541	tcagtcgataataggggtatggtgtgtttttgtcagagctcctgacaggccagccactg	600		
Db	541	TCAGTCGATATATGGGCTATTTGGTTGTGTTTTTCAGAGCTCCTGACAGCCAGCCACTG	600		
QY	601	atcccgagacataaatcaacttttaaaagttaacgggtttttccatggcagataacct	660		
Db	601	TGGCCTGGAAAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAAATTA	660		
QY	661	atcccgagacataaatcaacttttaaaagttaacgggtttttccatggcagataacct	720		
Db	661	ATCCCAAGACATCAATCAATCTTTAAAGTATACGGGTTTTTCCATGGCATCAGTATACCT	720		
QY	721	gagccagaacacatgaaactcttgaggaaaagtctcagatgttcatcctgtggtctgt	780		
Db	721	GAGCCAGACACATGGAACCTCTTGAGGAAAAGTTCTCAGATGTTTCATCCTGTGCTCTG	780		
QY	781	aacttcatgaaggggtgtctgaagatgaatccagatgacagataaacctgtgtcccaactc	840		
Db	781	AACTTCTGAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTACCTGTGTCCCAACTC	840		
QY	841	ctggagagctcctacttcttatttttaagagcccaattaaaaagaagaagcagtaat	900		
Db	841	CTGGAGAGCTCCTACTTTGATTCTTTTCAAGAGGCCCAATTTAAAGAAAAAGCACGTAAT	900		
QY	901	gaaggagaacacagagacgcacaggttaacttcgcgtcaaaagt	945		
Db	901	GAAGGAGAAAAACAGAACGCCAACAGGTACTTCCGCTCAAAAGT	945		

Qy	661	atccacagacatcaatcaatctttaaaagtaacgggtttttccatgcatacct	720
Db	661	ATCCCAAGACATCAATCAATCTTTAAAGTAACGGGTTTTTCCATGCATCAGTACT	720
Qy	721	gaqccagaagacatgaaactcttgagaaaagtctcagatgttcacctcgtggtctg	780
Db	721	GAGCCAGAAGACATGGAACCTCTTGAGAAAAGTTCTCAGATGTTCACTGTTGGCTCG	780
Qy	781	aactcatgaagggtgtctgaagataatccagatgcagattaaactgttcccaactc	840
Db	781	AACTTCATGAAGGGGTGCTGAAGATGAATCCAGATGCACAGATTAACCTGTTCCCACTC	840
Qy	841	ctggagagctcacttacttcttttcaagagggcccaattaaaaagaagacacgtat	900
Db	841	CTGGAGAGCTCCTACTTTGATTCTTTTCAAGAGGCCCAATTTAAAAGAAAGACCGTAAT	900
Qy	901	gaaggaagaaacagagacgccaacagt	929
Db	901	GAAGGAAGAAACAGAGACGCCAACAGGT	929
RESULT	3		
LOCUS	AX166534	1083 bp	DNA
DEFINITION	Sequence 25 from Patent WO0138503.		
ACCESSION	AX166534		
VERSION	AX166534.1	GI:14546879	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Plooman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R., Flanagan, P. and Clary, D.S.		
TITLE	Novel human protein kinases and protein kinase-like enzymes		
JOURNAL	Patent: WO 0138503-A 25 31-MAY-2001.		
FEATURES			
source	Location/Qualifiers		
	1..1083		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	366 a 199 c 225 g 293 t		
ORIGIN			
Query Match	91.3%;	Score 863;	DB 6; Length 1083;
Best Local Similarity	94.5%;	Pred. No. 5.9e-184;	
Matches 927;	Conservative	0; Mismatches	0; Indels 54; Gaps 1;
Qy	1	atggaaaagtataaaaattagctaaagactggaagggtcttatgggtgtatctcaaa	60
Db	1	ATGGAAGATATGAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGTGTATTTCAA	60
Qy	61	tcagaaacaaacctctggacaagttagctgttaaaaaattgtggaactcgaagat	120
Db	61	TCCAGAAACAAACCTCTGGCACTACTAGCTGTAAAAAATTTGTGGAATCTGAAGAT	120
Qy	121	gatcctgtgttaagaaaatagcactaagagaatactatgttggaagcaataaaacat	180
Db	121	GATCCTGTGTGTAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAACT	180
Qy	181	ccaaacttgtaacctcatcgaggtgttcagagaaaaagaaaatgcatttagttttt	240
Db	181	CCAAATCTTGTGAACCTCATCGAGGTGTTCCAGAGAAAAGAAAATGCATTTAGTTTTT	240
Qy	241	gaactactgtgatcatcaccttttaaatgagctggaagaaaccccaaatgaggttctgat	300
Db	241	GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTGTGAT	300
Qy	301	gagtgatcaaaagcgtattatggcaaaccttcaagctcttaattctgtcatatacat	360
Db	301	GAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTCTGTCTATATACAT	360
Qy	361	aactgtattcacagagatataaaacacgtgaaaaatattcttaataacgaagaaataatc	420
Db	361	AACGTATTTCACAGAGATATAAAACCTGAAAAATTTCTTAATAACTAAGCAAGGAATAATC	420
Qy	421	aagatttgtacttcgggttttgcaaaattctga	454
Db	421	AAGATTGTGACTTCGGGTTTGCACAATTTCTGAGTAATTCACATGTTGGGAGAATTGAC	480
Qy	455	-----ttccaggagatgcctacacacgattatgaact	486
Db	481	TTACCTGATCTGATTGATGCAATTTCCAGGAGATGSCCTACCCGGATTATGTAGCT	540
Qy	487	acgaatgttaccagagctcctgaactctcttctgtggagatactcagtagtattcttcagtc	546
Db	541	ACGATGTGTACCGAGCTCCTCAACTTCTTGTGGAGATACTCAGTATGTGTTCTTCAGTC	600
Qy	547	gatataatgagctatgtgtgttttttcagagagctcctgaagccagccactgtgacct	606
Db	601	GATATATGGCTATGTGTTGTTTTGCAGAGCTCTCTCAGAGGCCACCCACTGTGGCCT	660
Qy	607	ggaaaatcagatgtggaccaactttatctgataatcagacaactagggaaaattaatccca	666
Db	661	GGAAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGGAAAATTAATCCCA	720
Qy	667	agacatcaatcaactctttaaagtaacgggtttttccatggcatcagtataccctgagcca	726
Db	721	AGACATCAATCAATCTTTTAAAGTAACGGGTTTTTCCATGGCANTCAGTATACCTGAGCCA	780
Qy	727	gaagacatggaaaactcttgaggaaaagtctcagatgttctcagatgttctcgtggaacttc	786
Db	781	GAAGACATGGAACCTTTGAGGAAAAGTTCTCAGATGTTCACTCTGTGGCTCTGAACCTC	840
Qy	787	atgaagggtgtctgaagatgaatccagatgaacattacacgtttcccaactcctctgag	846
Db	841	ATGAAGGGGTGTCGAAGATGAATCCAGATGCACAGATTAACTGTGCCAACTCCTGGAG	900
Qy	847	agctcctactttgatcttttcaagaggcccaaatataaagaaaagacacgtaatgaagga	906
Db	901	AGCTCTACTTTGATTCTTTTCAAGAGGCCCAATTAAGAAAAGACGTAATGAAGGA	960
Qy	907	agaaacagaagacgcacaacag	927
Db	961	AGAAACAGAGACGCCAACAG	981
RESULT	4		
LOCUS	AX107720	1041 bp	DNA
DEFINITION	Sequence 9 from Patent WO0123579.		
ACCESSION	AX107720		
VERSION	AX107720.1	GI:13923201	
KEYWORDS	human.		
SOURCE	Homo sapiens		
REFERENCE			
AUTHORS	Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.		
TITLE	Human kinase proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0123579-A 9 05-APR-2001.		
FEATURES	Location/Qualifiers		
	1..1041		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	344 a 190 c 228 g 279 t		
ORIGIN			
Query Match	88.8%;	Score 839;	DB 6; Length 1041;

Best Local Similarity 90.8%; Pred. No. 1.4e-178;  
Matches 945; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 atggagaagatgaataaattagctaaagctggagaagggtcttatgggtgtgattcaaa 60  
|||||  
Db 1 ATGGAAAAGATGGAATAATTAGCTAAGACTGGGAAGGGCTTATGGGTGTGATTCAA 60  
61 tgcgagaacaaacctctggacaagtagtgcgtgttaaaaaatttggaaatcgaagat 120  
Db 61 TGCAGAAACAAACCTCTGGACAAGTAGTACGTGTAAAAAATTTGGAAATCTGAAGAT 120  
QY 121 gatcctgtgttaacgaataagcaataagagaatacgtatgttgaagcaatataaacat 180  
Db 121 GATCCTGTGTAAAGAAATAGACACTAAGAGAAATACGTATGTGTAAGCAATATAACAT 180  
QY 181 ccaaatcttggaaacctcgtgcaggtgttcaggagaacaaaggaaatgcatttattttt 240  
Db 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTTAGTTTTT 240  
QY 241 gaactctgtgatacaacttttaaatgagctgaaagaaacccaaatggaggttgcgtgat 300  
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300  
QY 301 ggagtgatacaaaagcgtattatggcaaaccttcaagctcttaatttctgtcatatacat 360  
Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360  
QY 361 aactgtattcacagagataaaacctgaaactgaaatattcctaactaaagcaaggaataac 420  
Db 361 AACTGTATTACAGAGATATAAACCTGAANAATATTCTAATACATAAGCAAGGAATAATC 420  
QY 421 aagatttgcacttcgggttgcacaaattctgattccagagagatgcctacacgattat 480  
Db 421 AAGATTGTGACTTCGGGTTTGCACAAATCTGATTTCAGAGAGATGCTACACCGATTAT 480  
QY 481 gtactacgagatgtaccgagctcctgaaactcttctgtggagatactcagatgatctct 540  
Db 481 GTAGCTAGGAGATGGTACCGAGCTCCTGAACCTTCTGTGGGAGATCTCAGTATGTTCT 540  
QY 541 tcagtcgataatggcctattgtgtgttttcagagactcctgcagagccagccactg 600  
Db 541 TCAGTCGATATATGGCTATGTGTTGTGTTTTCAGAGACTCCTGCAGAGCCAGCCACTG 600  
QY 601 tggcctgaaatacagatgtggacaaacttattctgtataatcagaacact- 650  
Db 601 TGGCCTGGAATAACAGATGTGGACCAACTTTATCTGTATTAATCAGAACACTAGTAGAGAG 650  
QY 651 ----- 650  
Db 661 GGGTTTCGCCATGTTGACCAGGCTGGTCTCGAACCTTCACGCTCAAGTGATCCACTGCC 720  
QY 651 -----aggaaaattaatcccaagacatacaatctttt 684  
Db 721 GTAGCCTCTCAAGCTGTGGAAATTACAGGAAAAATTAAATCCCAAGACATCAATCAATCTTT 780  
QY 685 aaagtaacgggttttttccatggcatcagtatacctgagccagaagacatgaaactctt 744  
Db 781 AAAAGTAACGGGTTTTTCATGGCATCAGTATACCTTGAGCCAGAGACATGGAACTCTT 840  
QY 745 gaggaaaagtctcagatgttcatcctgtggtctcgaacttcagactcatgaaggggtgtcgaag 804  
Db 841 GAGGAAAAGTCTCAGATGTTTCATCTCTGTGGCTCTGAACTTCATGAAGGGGTGTCTGAAG 900  
QY 805 atgaatccagatgacagattaaacctgttcccaactcctggagagctcctcactttgattct 864  
Db 901 ATGAATCCAGATGACAGATTAACTGTTCCCAACTCTCTGGAGAGCTCTTACTTTGATTCT 960  
QY 865 ttcaagagggcccaattaaagaaagcacgtaatgaaggaagaacagaagacgcgcaa 924  
Db 961 TTTCAAGAGGCCCCAAATTAAAGAAAGAACCCGTAATGAAGGAAGAAACAGAGACGCCAA 1020  
QY 925 caggtacttccgctcaaaagt 945  
|||||

Db 1021 CAGGTACTTCCGCTCAAAAGT 1041

RESULT 5

AX107714

LOCUS

AX107714

Sequence 3 from Patent WO0123579.

AX107714

VERSION

AX107714.1

GI:13923198

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.

TITLE

Human kinase proteins and polynucleotides encoding the same

JOURNAL

Patent: WO 0123579-A 3 05-APR-2001;

Lexicon Genetics Incorporated (US)

FEATURES

Location/Qualifiers

1..1068

Source

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 344 a 197 c 240 g 287 t

ORIGIN

Query Match 87.1%; Score 823; DB 6; Length 1068;

Best Local Similarity 90.6%; Pred. No. 5.7e-175;

Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 atggagaagatgaataaattagctaaagctggagaagggtcttatgggtgtgattcaaa 60

Db 1 ATGGAAAAGATGGAATAATTAGCTAAGACTGGGAAGGGCTTATGGGTGTGATTCAA 60

QY 61 tgcgagaacaaacctctggacaagtagtgcgtgttaaaaaatttggaaatcgaagat 120

Db 61 TGCAGAAACAAACCTCTGGACAAGTAGTACGTGTAAAAAATTTGGAAATCTGAAGAT 120

QY 121 gatcctgtgttaacgaataagcaataagagaatacgtatgttgaagcaatataaacat 180

Db 121 GATCCTGTGTAAAGAAATAGACACTAAGAGAAATACGTATGTGTAAGCAATATAACAT 180

QY 181 ccaaatcttggaaacctcgtgcaggtgttcaggagaacaaaggaaatgcatttattttt 240

Db 181 CCAAAATCTTGTGAACCTCATCGAGGTCTTCAGGAGAAAAAGGAAATGCATTTAGTTTTT 240

QY 241 gaatactgtgatacaacttttaaatgagctggaagaaacccaaatggaggttgcgtgat 300

Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAAGAAACCCAAATGGAGTTGCTGAT 300

QY 301 ggagtgatacaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360

Db 301 GGAGTGATCAAAAGCGTATTATGCAACACTTCAAGCTCTTAATTTCTGTCAATATACAT 360

QY 361 aactgtattcacagagataaaacctgaaacttcttaataacttaactaaagcaaggaataac 420

Db 361 AACTGTATTCACAGAGATATAAACCTGAAANATATCTAATAACTAAGCAAGGAATAATC 420

QY 421 aagatttgcacttcgggttgcacaaattctgatccagagagatgcctacacgattat 480

Db 421 AAGATTGTGACTTCGGGTTTGCACAAATCTCTGATTCAGAGAGATGCCTACACCGATTAT 480

QY 481 gtactacgagatgtaccgagctcctgaaacttctgtggagatactcaataggttct 540

Db 481 GTAGCTAGGAGATGGTACCGAGCTCCTGAACCTTCTGTGGGAGATCTCAGTATGTTCT 540

QY 541 tcagtcgataatggcctattgtgtgttttgcagagctcctgcagagccagccactg 600

Db 541 TCAGTCGATATATGGCTATGTGTTGTGTTTTCAGAGCTCTCTCAGAGCCAGCCACTG 600

QY 601 tggcctgaaatacagatgtggacaaacttattctgtataatcagaacact- 650

Db 601 TGGCCTGGAATAACAGATGTGGACCAACTTTATCTGTATTAATCAGAACACTAGTAGAGAG 650



|||||  
Db 601 TGCCGGGAAATCAGATGGGACCACTTTATCTGATAATCAGAACACTAGTAGAGCG 660  
Qy 651 ----- 650  
Db 661 GGCTTTCGCGCATGTTGACACAGGCTGGTCTCGAACTCTTGACGTCAAGTGATCCACCTGCC 720  
Qy 651 ----- 684  
Db 721 GTAGCCTCTCAAAAGTCTGGAATTAACAGGAAATTAATCCCAAGACATCAATCAATCTTT 780  
Qy 685 aaagttaacgggtttttccatgcacatcagatatacctgagccagagacatggaactctt 744  
Db 781 AAAGTAACGGGTTTTTCATGGCATCAGTATACCTGAGCCAGACATGGAACCTCT 840  
Qy 745 gaggaagaatttcagatgttccctgtggtctgaacttcgaacttcgaaggggtgctgaag 804  
Db 841 GAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACTTCATGAAGGGGTGCTGAAG 900  
Qy 805 atgaatccagatgacagattaaacctgttcccaactcctggagagctcctactttgattct 864  
Db 901 ATGATCCAGATGACAGATTAACTGTTCCTCAACTCTCGAGAGCTCTTACTTTGATTCT 960  
Qy 865 tttaagagggcccaataataaagaaagcacgtaataatgaaggaagaaacagaagacgcaa 924  
Db 961 TTTCAAGAGGCCCAATTAAGAAAGCAGCGTAATGAAGGAAGAAACAGAGAGCGCAA 1020  
Qy 925 caggt 929  
Db 1021 CAGGT 1025  
|||||

RESULT 6  
AX056404 911 bp DNA PAT 13-JAN-2001  
LOCUS AX056404 Sequence 48 from Patent WO0073469.  
DEFINITION AX056404  
ACCESSION AX056404  
VERSION AX056404.1 GI:12229111  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 911)  
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.  
TITLE Protein kinases  
JOURNAL Patent: WO 0073469-A 48 07-DEC-2000;  
Sugen, Inc. (US)  
FEATURES  
source  
1. .911  
/organism="Homo sapiens"  
BASE COUNT 293 a 182 c 178 g 258 t  
ORIGIN

Query Match 68.0%; Score 643; DB 6; Length 911;  
Best Local Similarity 100.0%; Pred. No. 1.6e-134; Mismatches 0; Indels 0; Gaps 0;  
Matches 643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 285 aaatggagttgctgagtgatcaaaagcgtattattatggcaaacacttcaagctcttaa 344  
Db 1 AAATGGAGTTGCTGATGGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAA 60  
Qy 345 ttctgctatatacaataactgtattcacagagatataaacctgaaatattcttaataac 404  
Db 61 TTTCTGTCATATACATACTGATTTCACAGAGATATAAACCTGAAATATCTTAATAAC 120  
Qy 405 taagcaaggaataataagatttgacttcgggttgacaaaattctgattccagagaa 464  
Db 121 TAAGCAAGGAATAATAAGATTGTGACTTCGGGTTTGACAAAATTCGTATTCAGAGAA 180  
Qy 465 tgcctacaccgattatgactacgagatggtacccgagctcctgaactcttctgtggaga 524

|||||  
Db 181 TGCCATACACCATATGATAGCTACGAGATGGTACCAGCTCTGAACCTTCTTGTGGGAGA 240  
Qy 525 tactcagtatggttcttcagtcgatatatgggctattggtgtgttcttcagagctcct 584  
Db 241 TACTCAGTATGGTTCTTTCAGTCGATATATGGGCTATTTGGTTGTGTTTTCAGAGCTCT 300  
Qy 585 gacagccagccactgtggcctggaaatcagatgtgaccacactttctctgataatcag 644  
Db 301 GACAGCCAGCCACTGTGGCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 360  
Qy 645 aacctaggaataataatcccaagacataatcaatctttaaaagtaacgggtttttcca 704  
Db 361 AACACTAGGAAATTAATCCCAAGACATCAATCTTTAAAGATTAACGGGTTTTTCCA 420  
Qy 705 tgcacatcagatatacctgagccagagacatggaactcttgaggaaaagtctcagatgt 764  
Db 421 TGGCATCAGTATACCTCAGCCACAGACATGAAACTCTTTGAGGAAAAGTTCTCAGATGT 480  
Qy 765 tcatctgtgctctgaacttcgaaggggtgctgaagatgaatccagatgacagatt 824  
Db 481 TCATCTCTGTGGCTCTGAACCTTCATGAAGGGGTGCTGAAGATGAATCCAGATGACAGATT 540  
Qy 825 aacctgttcccaactcctggagagctcctactttgattcttttcaagaggcccaataa 884  
Db 541 AACCTGTCCCAACTCTCGAGAGCTCTTACTTTGATTCTTTTCAAGAGGCCCAATTA 600  
Qy 885 aagaaaagcacgtaataatgaaggaagaaacagaagacgccaacag 927  
Db 601 AAGAAAAGCACGTAATGAAGGAAGAAACAGAGAGCGCCACAG 643  
|||||

RESULT 7  
AX056405 2615 bp DNA PAT 13-JAN-2001  
LOCUS AX056405 Sequence 49 from Patent WO0073469.  
DEFINITION AX056405  
ACCESSION AX056405  
VERSION AX056405.1 GI:12229112  
KEYWORDS Murinae gen. sp.  
SOURCE Murinae gen. sp.  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.  
REFERENCE 1 (bases 1 to 2615)  
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.  
TITLE Protein kinases  
JOURNAL Patent: WO 0073469-A 49 07-DEC-2000;  
Sugen, Inc. (US)  
FEATURES  
source  
1. .2615  
/organism="Murinae gen. sp."  
BASE COUNT 742 a 580 c 585 g 708 t  
ORIGIN

Query Match 59.2%; Score 559.2; DB 6; Length 2615;  
Best Local Similarity 81.8%; Pred. No. 1.1e-115; Mismatches 143; Indels 3; Gaps 1;  
Matches 658; Conservative 0; Mismatches 143; Indels 3; Gaps 1;  
Qy 135 gaaaatgactaagaagaatacgtatgttgaggcaattaaacacatcccaactctgtgaa 194  
Db 1 GAAATAGCCCTCGCGGAATTCGTTATGCTGAAG---TTGAACACCCCAACCTCGTGA 57  
Qy 195 cctcatcagaggtgttcagagaaaaaagaaatgaattgttttaactatcgtatca 254  
Db 58 CCTCATCGAGGTGTTCAGAAGAAAGAGAAAGATGCATCTAGTTTTTGACTACTGATCA 117  
Qy 255 tacacttttaaatgagctggaagaaaccccaaatggagttgctgtagagtgatcaaaag 314  
Db 118 CACACTGTTAAACGAGCTGGAGAGAAACCCCAACGGAGTTTCTGTATGGGTGATTAAG 177  
Qy 315 cgtattatggcaaacactcaagctcttaattctctcatatacataaactgtattccacag 374

```
Db 178 TGTCTATGCAAAACCTTCAAGCCCTTAACCTCTGTCTCACAAGCACAATGTATTCAFCG 237
QY 375 agataaaaaacctgaaaaattcttaataactaagcaaggaataatacaagatttggactt 434
Db 238 GGATGTAAAACCTTGAACACATCCTTAATACCAAGCAAGGGATGATTAAGATTGTGACTT 297
QY 435 cgggttggcaaaattctgattccaggagatgacctacacgattatgtagtaagatg 494
Db 298 TGGATTGGACGANTCTTAATTCAGGAGACGCCTACACAGACTATGTGGCCACCGGTG 357
QY 495 gtaccagctctgaactcttctgtggagatactcagtgattgttctcagtcgatatatg 554
Db 358 GTACCGAGCCCGAATCTCTCGTGGGAGACACGAAGTACGGTTCTCTGTAGACGTGTG 417
QY 555 ggcattgggttgttttggcagagctcctgacagggccagcccaactgtgacctggaatac 614
Db 418 GGCCGTGGCTGTGTTTGGCAGAGCTCCTGACGGGTACGCCACTCTGGCCGGGAAATC 477
QY 615 agatgggaccaactttatctgataatcagaacactaggaataataatcccaagacatca 674
Db 478 CGAGGTGGACCACTTTACCTGATCATCAGGACGTTGGGAAGCTGATTCCCAAGACACCA 537
QY 675 atcaacttttaaaagtaacggtttttccatgcacatcagttacacctgagccagagacat 734
Db 538 GTCTATCTTTAGAGTAACCAAGTTTTCGCGGCATCAGCATACCTTGACACAGAGGACAT 597
QY 735 ggaactcttggagaaagtctcagatgttccctgtgtctgtgacctggaacttcataagagg 794
Db 598 GGAGACTCTTTGAAGAAAAATTTCAAAATGTTCAAGCTGTGGCTTTAAGTTTCATGAAGGG 657
QY 795 gtctgaagatgaatccagatgacagatgaacctgttcccaactcctggagagctcta 854
Db 658 ATGCTTGAAGATGAATCTCTGATGAGAGGCTGACCTGTGCCAGCTGTGGACATGCCCTA 717
QY 855 ctttgattctttcaagagcccaaatcaaaagaaagcaactaatgaaggaagaaacag 914
Db 718 CTTTGTGCTTTTTCAGAGGATCAATGANAAGAAAGACCCGACGTGAGGGAGAACCG 777
QY 915 aagaagcccaacaggtacttcgct 938
Db 778 AAGGCCGACGAATACTCACTGCT 801

RESULT 8
AX107724
LOCUS AX107724 1819 bp DNA PAT 30-APR-2001
DEFINITION Sequence 13 from Patent WO0123579.
ACCESSION AX107724
VERSION AX107724.1 GI:13923203
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1819)
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and
Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 13 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..1819
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 539 a 368 c 406 g 505 t 1 others
ORIGIN
```

Query Match 53.7%; Score 507; DB 6; Length 1819;  
Best Local Similarity 69.8%; Pred. No. 5.8e-104;  
Matches 929; Conservative 0; Mismatches 0; Indels 402; Gaps 2;

```
QY 1 atggaataatgaaaaattagctaagactggagaaggggtcttatggggttgattcaaa 60
Db 138 ATGGAATAATGAAAAAATTAGCTAAGACTGGAGAAGGGTCTTTATGGGTTGTATTCAAA 197
QY 61 tgcagaaacaaacctctgacaagtagtagctgttaaaaaatttgggaactcagaat 120
Db 198 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAAAAATTTGTGGATCTGAAAGAT 257
QY 121 gatcctgttggtaagaaaaatagcactaagagaatacgtatgttgaagcaataaaacat 180
Db 258 GATCCTGTTGTTTAAAGAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTTAAAAACAT 317
QY 181 ccaaatcttggaaacctcatcagagggtgttcagggagaaaaaggaatacatttagtttt 240
Db 318 CCAAAATCTTGTGAACCTCATCGAGGTGTTCAGGAGAAAAAGGAAATGCATTTAGTTTTT 377
QY 241 gaatactgtatcatcacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300
Db 378 GAATACTGTGATCATACACITTTTAAATGAGCTGGAAAGAACCCAAATGGAGTTGCTGAT 437
QY 301 ggagtgtacaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
Db 438 GGAGTGATCAAAAGCGTATTATGGCAAAACACTTCAAGCTCTTAATTTCTGTCTATACAT 497
QY 361 aactgtattcacagagatataaaacctgaaaatatcttaataactaagaagaataatc 420
Db 498 AACTGTATTACAGAGATATAAAACCTGAAATATTCTAATACTAAGCAAGGAATATC 557
QY 421 agagttctgactcgggtttgcacaaattctg----- 453
Db 558 AAGATTTGTGACTTCGGGTTTGCACAAATCTGAGTTGGACTTTCATCTTCTCTGCTGCC 617
QY 454 ----- 453
Db 618 TCCTTGATTGCTTAATAGTTGACCTTCTGAATTCITTTTCTGCCAATTCAGAGATTTT 677
QY 454 ----- 453
Db 678 CTCCTGGCTGGATCCATTGCTGACACAGTGTTCACCATGGGGCCAGGCTCATCTCGA 737
QY 454 ----- 453
Db 738 ACTTCTGGCTCAAGTGATCTTCCACCTCGGCTCCCAAAAGTCTGGATTGCAAGTGTG 797
QY 454 ----- 453
Db 798 AGCCACCGTCCCGACGACAGATTTTCAAAACAATACTACTGAGAGCTCAACAAGATGT 857
QY 454 ----- 474
Db 858 TTAGTGGGAACACAAATTTTGAACAAATTTCTTGAGAACGCATTCAGGAGATGCCTACACC 917
QY 475 gattatgactacagatggtaccagctcctgaactctgttgggagatactcagtat 534
Db 918 GATTATGTAGTACAGAGATGTTACCGAGCTCCTGAACCTCTGTGGGAGATACTCAGTAT 977
QY 535 ggtcttcagtcgatatatgggtatgtgttgggttcagagctcctctgacagggcag 594
Db 978 GGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGAGCTCTGACAGGCCAG 1037
QY 595 ccactgtggcctggaataatcagatgtggaccaactttatctgataatcagaacct---- 650
Db 1038 CCACCTGGCCCTGGAANAATCAGATGTGGACCAACTTTATCTGATTAATCAGAACACTAGTA 1097
QY 651 ----- 650
Db 1098 GAGACGGGTTTCGCCATGTTGACACGCGTGTCTCGAACTCTTGACGTCAAGTGATCCA 1157
QY 651 ----- 678
Db 1158 CCTGCCGTAGCTCTCAAAAGTGTGGAATTACAGGAAAAATTAATCCCAAGACATCAATCA 1217
QY 679 atctttaaagtaacgggtttttccatggcatcagtatacctgagccagagacatggaa 738
```

||||| 1218 ATCTTTAAAGTACGGGTTTTTCATGGCATCAGTATACCTGAGCCAGACATGGAA 1277

Qy 739 actcttgaggaaagtcttcagatgttcactcctgtgctctgaacttcataagagggtgt 798

Db 1278 ACTCTTGAGGAAAGTCTCAGATGTTTCATCCTGTGGCTGCTGAACTTCAATGAAGGGTGT 1337

Qy 799 ctgaagatgaatcacatgacagattaacctgttcccaactcctggagagctcctacttt 858

Db 1338 CTGAAGATGAATCCAGATGACAGATTAACCTGTTCCTCAACTCCTGGAGAGCTCTACTTTT 1397

Qy 859 gattcttttcaagggcccaattaaagaaaagcacgtaataagagaagaaacacagaaga 918

Db 1398 GATTCTTTTCAAGAGGCCCAATTAAGAAAGAACGACGTAATGAAGAGAAACACAGAAGA 1457

Qy 919 cgccaacaggt 929

Db 1458 CGCCAACAGGT 1468

RESULT 9

AX107712 561 bp DNA PAT 30-APR-2001

LOCUS Sequence 1 from Patent WO0123579.

DEFINITION AX107712

ACCESSION AX107712

VERSION AX107712.1 GI:13923197

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 561)

AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.

TITLE Human kinase proteins and polynucleotides encoding the same

JOURNAL Patent: WO 0123579-A 1 05-APR-2001;

Lexicon Genetics Incorporated (US)

FEATURES

source 1..561

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 189 a 87 c 112 g 173 t

ORIGIN

Query Match 48.1%; Score 454.4; DB 6; Length 561;

Best Local Similarity 99.8%; Pred. No. 3.8e-92;

Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atggaaaagtataaaaaattagctaaagctgagagagggtcttatggggtgtattcaaa 60

Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAAA 60

Qy 61 tcgagaaacaaacctctggacaagtagctgttaaaaaatttggaaatctgaagat 120

Db 61 TCGAGAAACAAACCTCTGGACAAGTAGTGTGTTAAAAAATTGTGGAATCTGAAGAT 120

Qy 121 gatcctgttgaagaaatagcactaagagaatacgtatgttgaagcaattaaacat 180

Db 121 GATCCTGTGTTAAGAAAAATAGCCTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180

Qy 181 ccaaatcttgaacctcatcgaggtgttcaggagaaaaaggaataatgcatttatttt 240

Db 181 CCNAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAAAGAAAAATGCAATTTAGTTTTT 240

Qy 241 gaatactgtatatacacttttaaatgagctgagaaagaaaccccaaatgaggtgctgat 300

Db 241 GAATACTGTATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGAGGTGCTGAT 300

Qy 301 ggaagtatacaaacgctattatggcaaacacttcaagctcttaattcttgcataacat 360

Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTTCTGTCTATATACAT 360

Qy 361 aactgtattcacagagataaaaaaccctgaaataattcttaataactaagcaaggaataac 420

Db 361 AACTGTATTACAGAGATATAAAACCTGAAATAATTCTAANTAATAAGCAAGGAATAATC 420

Qy 421 aagatttgcacttcgggttttgcacaaattctgatt 456

Db 421 AAGATTTCTGACTTCGGGTTTGCACAAAATTTCTGAGT 456

RESULT 11

HSSTHPKB

Qy 361 aactgtattcacagagataaaaaaccctgaaataattcttaataactaagcaaggaataac 420

Db 361 AACTGTATTACAGAGATATAAAACCTGAAATAATTCTAANTAATAAGCAAGGAATAATC 420

Qy 421 aagatttgcacttcgggttttgcacaaattctgatt 456

Db 421 AAGATTTCTGACTTCGGGTTTGCACAAAATTTCTGAGT 456

RESULT 10

AX107718 594 bp DNA PAT 30-APR-2001

LOCUS Sequence 7 from Patent WO0123579.

DEFINITION AX107718

ACCESSION AX107718

VERSION AX107718.1 GI:13923200

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 594)

AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.

TITLE Human kinase proteins and polynucleotides encoding the same

JOURNAL Patent: WO 0123579-A 7 05-APR-2001;

Lexicon Genetics Incorporated (US)

FEATURES

source 1..594

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 203 a 94 c 115 g 182 t

ORIGIN

Query Match 48.1%; Score 454.4; DB 6; Length 594;

Best Local Similarity 99.8%; Pred. No. 3.8e-92;

Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atggaaaaglatgaaaaattagctaaagctgagagagggtcttatgggtgtattcaaa 60

Db 1 ATGGAAAAGTATGAAAAATTAGCTAAGACTGGAGAGGGTCTTATGGGGTGTATTCAAA 60

Qy 61 tcgagaaacaaacctctggacaagtagctgttaaaaaatttggaaatctgaagat 120

Db 61 TCGAGAAACAAACCTCTGGACAAGTAGTGTGTTAAAAAATTGTGGAATCTGAAGAT 120

Qy 121 gatcctgttgaagaaatagcactaagagaatacgtatgttgaagcaattaaacat 180

Db 121 GATCCTGTGTTAAGAAAAATAGCCTAAGAGAAATACGTATGTTGAAGCAATTAAACAT 180

Qy 181 ccaaatcttgaacctcatcgaggtgttcaggagaaaaaggaataatgcatttatttt 240

Db 181 CCNAATCTTGTGAACCTCATCGAGGTGTTCCAGGAGAAAAAGAAAAATGCAATTTAGTTTTT 240

Qy 241 gaatactgtatatacacttttaaatgagctgagaaagaaaccccaaatgaggtgctgat 300

Db 241 GAATACTGTATACACTTTTAAATGAGCTGGAAAGAAACCCCAATGAGGTGCTGAT 300

Qy 301 ggaagtatacaaacgctattatggcaaacacttcaagctcttaattcttgcataacat 360

Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACACTTCAAGCTCTTAATTTCTGTCTATATACAT 360

Qy 361 aactgtattcacagagataaaaaaccctgaaataattcttaataactaagcaaggaataac 420

Db 361 AACTGTATTACAGAGATATAAAACCTGAAATAATTCTAANTAATAAGCAAGGAATAATC 420

Qy 421 aagatttgcacttcgggttttgcacaaattctgatt 456

Db 421 AAGATTTCTGACTTCGGGTTTGCACAAAATTTCTGAGT 456

LOCUS	HS5THPKB	1363 bp	mRNA	PRI	06-FEB-1997
DEFINITION	H.sapiens mRNA KIAALRE for serine/threonine protein kinase.				
ACCESSION	X66358				
VERSION	X66358.1	GI:36614			
KEYWORDS	serine/threonine protein kinase.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Meyerson,M.L.				
TITLE	Submitted (12-MAY-1992) M.L. Meyerson, Massachusetts General Hospital, Cancer Center, Bldg 149, 13th Street, Charlestown MA 02129, USA				
JOURNAL	2 (bases 1 to 1363)				
REFERENCE	Meyerson,M., Enders,G.H., Wu,C.L., Su,L.K., Gorka,C., Nelson,C., Harlow,E. and Tsai,L.H.				
AUTHORS	A family of human cdc2-related protein kinases				
TITLE	EMBO J. 11 (8), 2909-2917 (1992)				
JOURNAL	92347325				
MEDLINE					
FEATURES	Location/Qualifiers				
source	1..1363				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/cell_type="HeLa"				
	/cell_line="HeLa"				
	/clone_lib="HeLa"				
CDS	Join(214..669,X66359.1:1..54,670..1290)				
	/note="KIAALRE coding sequence with 54 bp insertion"				
	/product="serine/threonine protein kinase"				
	/db_xref="PID:e49306"				
	214..1290				
CDS	/codon_start=1				
	/label="sthpkin_CDS"				
	/product="serine/threonine protein kinase"				
	/db_xref="PID:e49306"				
	214..1290				
	/codon_start=1				
	/product="serine/threonine protein kinase"				
	/protein_id="CAA7002.1"				
	/db_xref="GI:36615"				
	/translation="MMBEYKGIKIGEGSVVFKQRNEDTCQIVAKKFLSESDVPIK KIAALREMLKQHPNLVNLSEVFKRRRLHLFYCDHTVLHDLDRVQGVPHDLYVATITWOTLQAVNFCNKHNCITRDVKENILKHSVILKCDQFARLLTGTPSDYITDVATYKSPPELLVGTGTGPPVDVWRAIGCVFAELLGSLVPLWPKGSVDQYLIRKTLGLDILPHQOVFSTNGITFSVCKIPDEMEPLKFPNII SYPALGLLKGCLHMDPTERRLTCEQLHRRYFIREILEDALAKEDHPKTRKLRKSRHHCFTETSKLOYLPQLTGSSILPALNDKKYYCDTKLNLRFNPI"				
BASE COUNT	412 a	308 c	309 g	334 t	
ORIGIN					
	Query Match 43.8%; Score 413.8; DB 9; Length 1363;				
	Best Local Similarity 66.6%; Pred. No. 5.1e-83;				
	Matches 608; Conservative 0; Mismatches 302; Indels 3; Gaps 1;				
QY	1	atggaaagatgaaataatagctaagctaggagaaggtcttatgggttgattcaaa	60		
Db	217	ATGGAGAAGTATGAAAAAATGGGAAAATGGAGAAGGATCCTATGGAGTTGTTTCAA	276		
QY	61	tgcagaacaaacctctggacaagttagctgtgtataaaatttgtgaatctgaagat	120		
Db	277	TGTAGAAAACGGGACACGGGTCCAGATTGTGGCCATCAGAAAGTTCTTGGAAATCAAGAA	336		
QY	121	gattctgtgttaagaaataagcactaagagaataatgaagtgtgaagcaataataacat	180		
Db	337	GACCTGTCTAAGAAGAAATGCCCTTCGGGAAATCCCAATGCTCAAGCAACTCAAGCAT	396		
QY	181	ccaatatctgtgaacctcatcgaggtgttcaggagaaaaaggaaatgcatttagtttt	240		
Db	397	CCCAACCTTGTAACCTCCTCGAAGTCTTCAGAGGAARACGGAGCGTTCACCTGGTGTTT	456		
QY	241	gaatactgtgatcatcaccttttaaatgagctggaagaaacccaataggattgctgat	300		

FEATURES	Location/Qualifiers
source	1..3153 /organism="Oryctolagus cuniculus" /db_xref="taxon:9986" /sex="male" /tissue_type="brain" 322..2022 /note="corresponds to mouse KKIAMRE-beta" /codon_start=1 /product="Ser/Thr kinase KKIAMRE" /protein_id="BAA88508.1" /db_xref="GI:6624279"
CDS	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

Matches 508; Conservative 0; Mismatches 352; Indels 3; Gaps 1;			
QY	1	atggaagaatgataaaattagctgaagactggagaaggtcttatgtgggtgtattcaca	60
Db	376	ATGGA AAAATATGAAAACCTGGTGGTTGGTGGAGAAGGAGTTATGGAATGCTGATGAAG	435
QY	61	tgcagaaacaaacacctctggacaagtagctgttataaaatttggaaatctgaagat	120
Db	436	TGTAGGAATAAAGATACTGGAGAATGTCGCCATTAAGAAGTCTTAGAAGTACGAT	495
QY	121	gactcgtgtttaaagaataagcactaagagaataacgtatgttgaagcaaatataaacat	180
Db	496	GACAAAATGTTAAAGAGATGCAATGCGAGAAATCAAGTTACTAAAGCAACTTAGGCAT	555
QY	181	ccaaactcttgaaacctcagcaggtgttcagagagaaaaaagaaatacatttagtttt	240
Db	556	GAAAACCTTGCTGAATCTCTTGGAGTGTTGAAGAAAAAAGCATGTGCTAGCTCTTT	615
QY	241	gaatactgtgatacacacttttaaatagctggagaagaaacccaaatggagttgctgat	300
Db	616	GAATTTGTGACCAACACAAATCTTGATGACTTGGAGCTCTTTCCAAATGGACTAGACTAC	675
QY	301	ggagtgcataaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat	360
Db	676	CAAGTAGTTCAAAAGTATTGTTTCAGATTATTAATGGAATGGATTGTCACAGTCAC	735
QY	361	aactgtattccagagataaaacctgaataattcttaataactaagcaaggaataatc	420
Db	736	AATATCATACACAGATATAAGCCAGACAGATATATTAGTCTCCAGTGTGGCGTTGTC	795
QY	421	aagatttgtactcgggtttgcacaaattctg---attccaggagatgcctacaccgat	477
Db	796	AGCTATGCGAATTTGGATTTCGCGAACATTTGCGAGCTCTCTGGGGAGTTTATACTGAT	855
QY	478	tatgtactacagatggtaccagactcctgaactcttctgtggagatactcagtaggt	537
Db	856	TATGTGGCAACCCGATGTCACAGACTCCAGACTATTGGTTGGTGAATGCAAGTAGGC	915
QY	538	tcttcagtcataatagggctatggtgtgttttgcagagctcctgcagagccagcca	597
Db	916	AGGCTGTGTATGTGGGCCATTTGGTGTCTGGTAACGTGAATGTTATGGGGAAACC	975
QY	598	ctgtggacctggaatacagatgtggaccactttatctgataatcagaacactaggaaa	657
Db	976	CTAATTTCTCGAGATTCTGATATTGATCAGCTATATATCATATATTGATGCTTTAGTGAAT	1035
QY	658	ttaatcccaagacatcaatcaatctttaaaagtaacgggtttttccatggcatcagata	717
Db	1036	CTAATTCGAAGCATCAGGAGCTTTTATATAAATCTCTGTGTTGCTGGAGTAAGGTG	1095
QY	718	ctgagccagaagacatggaactcttgaggaaagttctcagatgttctatcctgtgct	777
Db	1096	CCTGAATATCAAGGAAGAGAACCTCTTGAAGACGCTATCTTAAGCTCTCTCAAGTGGTG	1155
QY	778	ctgaactctatgaaggggtgctgaagatgaatccagatgacagataacctgttccaa	837
Db	1156	ATAGATTTCAGAAAGAAATGCTTACATATTGACCCCGAAAAGACCCCTTCTGTGCTGAG	1215
QY	838	ctcctgagagctcctactttga 860	
Db	1216	CTCTACACCATGATTTCTTTCA 1238	
RESULT 14			
AB029067			
LOCUS			
DEFINITION	AB029067	1518 bp	mRNA
ACCESSION	AB029067		
VERSION	AB029067.1	GI:6594327	
KEYWORDS	Ser/Thr kinase KIAMRE-delta.		
SOURCE	Mus musculus (strain:C57BL/6) adult brain tissue		
			Uni-ZAP cDNA to mRNA.

ORGANISM	Mus musculus		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae; Mus. 1 (sites)		
AUTHORS	Sassa,T., Gomi,H., Sun,W., Ikeda,T., Thompson,R.F. and Itohara,S.		
TITLE	The Murine KIAMRE gene: variants, dual promoters, expression and chromosomal localization		
JOURNAL	Unpublished (1999)		
REFERENCE	2 (bases 1 to 1518)		
AUTHORS	Sassa,T., Gomi,H. and Itohara,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUN-1999) to the DDBJ/EMBL/GenBank databases.		
	Takayuki Sassa, The Institute of Physical and Chemical Research (RIKEN), Brain Science Institute (BSI): 2-1 Hirosawa, Wako, Saitama 351-0198, Japan [E-mail: sassa@brain.riken.go.jp, Tel:81-48-467-9724(ex.7716), Fax:81-48-467-9725]		
FEATURES	Location/Qualifiers		
source	1..1518		
	/organism="Mus musculus"		
	/strain="C57BL/6"		
	/db_xref="taxon:10090"		
	/dev_stage="adult"		
	/tissue_lib="Uni-ZAP"		
gene	/tissue_type="brain"		
	530..1518		
	/gene="Kkm"		
CDS	530..>1518		
	/gene="Kkm"		
	/note="cdc2-related kinase"		
	/codon_start=1		
	/product="Ser/Thr kinase KIAMRE-delta"		
	/protein_id="BAA88429.1"		
	/db_xref="GI:6594328"		
	/translation="MEKYENLGLVCEGSGYGMVKCRNKDSRIVAIAKKFFLESDDKKMVKIAMREIKLLKOLRHEINLVNLLRKKKRNLYLVEFVDHTLDDLLKLPNGLDYOVVQYLFQIINGIGFCHSHNIHROI KFNELVLSQSGVVKLDFGFTLAAPGEVYDVYATFWRAPELLGVDYKTKRAVDIWAIGCLVIEMLGPDLFPGESIDQLHIMTCLGNLIPRQELFYKNPVFAGVRPEVKDAEAPLESRYPLKPEAVISLAKKLIHIDPDKRPFCDLLIRHDFPDGFAERFSQELQLKLEKDARNNSLPPKSKKKKKKKK"		
BASE COUNT	421 a 303 c 382 g 412 t		
ORIGIN			
Query Match	28.8%; Score 272.6; DB 10; Length 1518;		
Best Local Similarity	58.8%; Pred. No. 2.7e-51;		
Matches	511; Conservative 0; Mismatches 349; Indels 9; Gaps 2;		
QY	1	atggaagaatgataaaattagctgaagactggagaaggtcttatgtgggtgtattcaca	60
Db	530	ATGGAGAATGACGAGAACCTTAGGATGTTGGTGGAGAGGGAGTTACGGAATGTTGATGAAG	589
QY	61	tgcagaaacaaacacctctggacaagtagctgttataaaatttggaaatctgaagat	120
Db	590	TGTAGGAACAAAGACAGTGGAGAATTTGGCCATTAAAGAGTTCTTAGAAGTTCCTTAGAAGTGACGAT	649
QY	121	gactcgtgtttaaagaataagcactaagagaataacgtatgttgaagcaaatataaacat	180
Db	650	GATAAATGTTTAAAGAAGATCGCAATGCGAGAGATCAAGTTACTAAAGCAACTCGCGCAT	709
QY	181	ccaaatcttgigaacctcatcagggtgttcaggagaaaaaaggaaatcatttagtttt	240
Db	710	GAAAATTTGGTGAATCTGTTGGAAGTATGTAAGAAAAAAGACCTTGGTACTAGTCTTT	769
QY	241	gaatactgtgatacacacttttaaatagctggagaagaaacccaaatggagttgctgat	300
Db	770	GAAATTTGTGACACACAGATTTCTTGATGACTTGAAGACTCTTTCCAAATGGACTAGACTAT	829
QY	301	ggagtgcataaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat	360
Db	830	CAAGTAGTTCAAAAATATTGTTTTCAGATTATTAATGGAATGGGATTTGCCACAGTCAC	889
QY	361	aactgtattccagagataaaacctgaaatattctataactaagcaaggaataatc	420



```

Db 960 CCTGAGTCAAGATGCAAGAGCAGAGCCTCTTGAGAGCCGCTATCCTAAGCTCCCTGAA 1019
QY 772 gtggctctgaacttcataaggggtgctgaagatgaatccagatgacagattaaacctgt 831
Db 1020 GCTGTGATCAGTTAGCAAGAAATGCTTGACATTTGACCCAGACAAAAGGCCCTTCTGT 1079
QY 832 tcccaactctggagagctcctacttga 860
Db 1080 GCTGACCTTCTGGCCATGACTTCTTCA 1108

```

Search completed: May 2, 2002, 16:29:40  
Job time: 9562 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 2, 2002, 17:13:01 ; Search time 234.75 Seconds  
(without alignments)  
3451.213 Million cell updates/sec

Title: US-09-671-050-11  
Perfect score: 945  
Sequence: 1 atgggaaagtgaataatt.....aggctactccgctcaaaagt 945

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues  
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_1101.\*  
1: /SIDSL1/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	945	100.0	945	22	Human kinase cDNA
2	929	98.3	972	22	Human kinase cDNA
3	863	91.3	1083	22	Polynucleotide seq
4	839	88.8	1041	22	Human kinase cDNA
5	823	87.1	1068	22	Human kinase cDNA
6	843	88.0	911	22	Novel protein kina
7	559.2	59.2	2615	22	Novel protein kina
8	507	53.7	1819	22	Human kinase cDNA
9	454.4	48.1	561	22	Human kinase cDNA
10	454.4	48.1	594	22	Human kinase cDNA
11	287	30.4	1698	21	Rabbit KKIAMRE kin

12	287	30.4	3080	21	AAA29745	Rabbit KKIAMRE kin
13	227.4	24.1	1513	22	AAZ51208	Human lost in leuk
14	227.4	24.1	1773	22	AAZ506724	Polynucleotide seq
15	211.4	22.4	1667	21	AAZ51207	Rat lost in leukae
16	211.4	22.4	1738	21	AAZ51206	Rat lost in leukae
17	175.8	18.6	1866	22	AAF44672	Novel protein kina
18	171.8	18.2	903	20	AAZ07476	Pneumocystis carin
19	163.4	17.3	1476	21	AAZ37835	Cylin-dependent k
20	163.4	17.3	1635	19	AAV71073	Green fluorescent
21	163.4	17.3	1635	19	AAV71074	Human prostate can
22	161.4	17.1	1322	21	AAZ15889	Human polynucleoti
23	136.6	14.5	1825	22	AAZ58432	Cyclin dependent p
24	134.6	14.2	1089	15	AAZ063490	CDK1 protein. Can
25	131	13.9	1002	18	AAZ64448	Candida gene encod
26	131	13.9	1002	18	AAZ08264	Human ORFX ORF1486
27	128.2	13.6	6015	21	AAZ75931	Arabidopsis thalia
28	128	13.5	1310	21	AAZ35344	Novel protein kina
29	126.6	13.4	1899	22	AAZ44678	Arabidopsis thalia
30	126.4	13.4	1307	21	AAZ51531	Arabidopsis thalia
31	126	13.3	1293	21	AAZ36413	Human polynucleoti
32	125.6	13.3	1788	22	AAZ60218	Zea mays DNA fragm
33	125.2	13.2	1326	21	AAZ43758	Polynucleotide D1
34	117.6	12.4	936	22	AAZ58252	Polynucleotide D1
35	117.6	12.4	936	22	AAZ58254	Polynucleotide D1
36	117.6	12.4	936	22	AAZ58257	Polynucleotide D2
37	117.6	12.4	936	22	AAZ58259	Polynucleotide D2
38	117.6	12.4	936	22	AAZ58255	Polynucleotide D1
39	117.6	12.4	936	22	AAZ58252	Polynucleotide D1
40	117.2	12.4	936	22	AAZ58254	Polynucleotide D1
41	117.2	12.4	936	22	AAZ58257	Polynucleotide D2
42	117.2	12.4	936	22	AAZ58259	Polynucleotide D2
43	117.2	12.4	936	22	AAZ58255	Polynucleotide D1
44	117.2	12.4	936	22	AAZ58252	Polynucleotide D1
45	117.2	12.4	938	22	AAZ58255	Polynucleotide D2

ALIGNMENTS

RESULT 1  
AAD03817  
ID AAD03817 standard; cDNA: 945 BP.  
XX  
AC AAD03817;  
XX  
DT 19-JUN-2001 (first entry)  
XX  
DE Human kinase cDNA #6.  
XX  
KW Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..945  
FT /tag= a  
FT /product= "Human kinase #6"  
FT /note= "The coding region does not include stop codon"  
FT /partial  
XX  
XX WO200123579-A1.  
XX  
XX 05-APR-2001.  
XX  
XX 27-SEP-2000; 2000WO-US26621.  
XX  
XX 28-SEP-1999; 99US-0156511.  
XX  
XX (LEXI-) LEXICON GENETICS INC.  
XX  
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
XX

DR WPI: 2001-266166/27.  
 XX P-PSDB; AAE00495.  
 PT New isolated human kinase polynucleotide useful for generating  
 PT antibodies, as reagents in diagnostic assays and for screening for  
 PT compounds useful for treating mental, biological or medical diseases  
 XX  
 PS Claim 3; Page 33; 38pp; English.  
 XX  
 CC The present sequence is a cDNA encoding novel human protein  
 CC (NHP) known as human kinase. The human kinases share structural  
 CC similarity with animal kinases, more particularly serine or  
 CC threonine protein kinases. Human kinase cDNA is useful for the  
 CC detection of mutant human kinase for the diagnosis of disease,  
 CC and also as a therapeutic. It is useful for screening drugs  
 CC effective in the treatment of symptomatic or phenotypic  
 CC manifestations perturbing the normal function of NHP in the  
 CC body. The NHP nucleotide sequences are useful for generation of  
 CC antibodies, as reagents in diagnostic assays, for the  
 CC identification of other cellular gene products related to human  
 CC kinases, and as reagents in assays for screening compounds that  
 CC are useful for treating mental, biological or medical disorders.  
 CC NHP oligonucleotides are used as probes. The labelled NHP probes  
 CC are useful for screening human genomic library for identifying  
 CC polymorphisms and as primers in amplification assays to detect  
 CC mutations within the exons, introns and splice sites that can  
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
 CC encoding NHP products are used to genetically engineer cells  
 CC in vivo that functions as bioreactors in the body delivering a  
 CC continuous supply of NHP to the body. Nucleotide constructs  
 CC encoding functional NHPs are used in gene therapy for the  
 CC modulation of NHP expression.  
 XX  
 SQ Sequence 945 BP; 323 A; 165 C; 202 G; 255 T; 0 other;

Query Match 100.0%; Score 945; DB 22; Length 945;  
 Best Local Similarity 100.0%; Pred No. 4.4e-245;  
 Matches 945; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 atggaaagatgataaaattagctgaagctgagaggggtctattggttattcaaa 60  
 Db |||||  
 QY 1 atggaaagatgataaaattagctgaagctgagaggggtctattggttattcaaa 60  
 Db |||||  
 QY 61 tgcagaacaaacctctggacaagtagtgcgtttaaaaaattgtgaactcgaagat 120  
 Db |||||  
 QY 61 tgcagaacaaacctctggacaagtagtgcgtttaaaaaattgtggaaacctggaagat 120  
 Db |||||  
 QY 121 gatcctgtgttaagaaatagcactaagagaatacgtatgttggaagcaattaaacat 180  
 Db |||||  
 QY 121 gatcctgtgttaagaaatagcactaagagaatacgtatgttggaagcaattaaacat 180  
 Db |||||  
 QY 181 ccaaatctgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgattagtttt 240  
 Db |||||  
 QY 181 ccaaatctgtgaacctcatcgagggtgttcaggagaaaaaggaaaaatgattagtttt 240  
 Db |||||  
 QY 241 gaatactgtgatacacttttaaatgagctggaagaaacccaatggaggtgctgat 300  
 Db |||||  
 QY 241 gaatactgtgatacacttttaaatgagctggaagaaacccaatggaggtgctgat 300  
 Db |||||  
 QY 301 ggagtgatacaaacggtattatggcaaacacttcaagctttaattctgtcatacat 360  
 Db |||||  
 QY 301 ggagtgatacaaacggtattatggcaaacacttcaagctttaattctgtcatacat 360  
 Db |||||  
 QY 361 aactgtattcagagatataaaacctgaaataattcttaactaagcaaggaataac 420  
 Db |||||  
 QY 361 aactgtattcagagatataaaacctgaaataattcttaactaagcaaggaataac 420  
 Db |||||  
 QY 421 aagattgtgacttcgggtttgcacaaattctgattccaggagatgctcacacgattat 480  
 Db |||||  
 QY 421 aagattgtgacttcgggtttgcacaaattctgattccaggagatgctcacacgattat 480  
 Db |||||  
 QY 481 gtatctacagatggtaccagagctcctgaactcttctgtgggagatactcagtagttct 540

Db |||||  
 QY 481 gtatctacagatggtaccagagctcctgaactcttctgtggagatactcagtagttct 540  
 Db |||||  
 QY 541 tcaatgatataatgggtctattgtgtgtttttgcagagctcctgacagccagcactg 600  
 Db |||||  
 QY 541 tcaatgatataatgggtctattgtgtgtttttgcagagctcctgacagccagcactg 600  
 Db |||||  
 QY 601 tggcctgaaaaatcagatggtgacacactttatctgataatcagaaactagaaaaatta 660  
 Db |||||  
 QY 601 tggcctgaaaaatcagatggtgacacactttatctgataatcagaaactagaaaaatta 660  
 Db |||||  
 QY 661 atcccaagacatacaatcaatctttaaagaagtaacgggtttttccatggcagatacact 720  
 Db |||||  
 QY 661 atcccaagacatacaatcaatctttaaagaagtaacgggtttttccatggcagatacact 720  
 Db |||||  
 QY 721 gagccagaagacatggaactcttgagaaaaattctcagatgttcattctcgtgctctg 780  
 Db |||||  
 QY 721 gagccagaagacatggaactcttgagaaaaattctcagatgttcattctcgtgctctg 780  
 Db |||||  
 QY 781 aacttcaggaaggggtctgaagatgaatccagatgacagattaaacctgttcccaactc 840  
 Db |||||  
 QY 781 aacttcaggaaggggtctgaagatgaatccagatgacagattaaacctgttcccaactc 840  
 Db |||||  
 QY 841 ctggagagctcctactcttcttcaagagcccaaaatgaaagaaagacagtaaat 900  
 Db |||||  
 QY 841 ctggagagctcctactcttcttcaagagcccaaaatgaaagaaagacagtaaat 900  
 Db |||||  
 QY 901 gaaggaagaacacagaagcgcacacaggtacttccggtcaaaagt 945  
 Db |||||  
 QY 901 gaaggaagaacacagaagcgcacacaggtacttccggtcaaaagt 945

RESULT 2  
 AAD03814  
 ID AAD03814 standard; cDNA: 972 BP.  
 XX  
 AC AAD03814;  
 XX  
 DT 19-JUN-2001 (first entry)  
 XX  
 DE Human kinase cDNA #3.  
 XX  
 KW Human; kinase; gene therapy; bioreactor; mental disorder;  
 KW biological disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..972  
 FT /\*tag= a  
 FT /product= "Human kinase #3"  
 FT /note= "The coding region does not include stop codon"  
 FT /partial  
 XX  
 PN WO200123579-A1.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 27-SEP-2000; 2000WO-US26621.  
 XX  
 PR 28-SEP-1999; 99US-0156511.  
 XX  
 PA (LEXI-) LEXICON GENETICS INC.  
 XX  
 PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
 XX WPI: 2001-266166/27.  
 DR P-PSDB; AAE00492.  
 XX  
 PT New isolated human kinase polynucleotide useful for generating  
 PT antibodies, as reagents in diagnostic assays and for screening for  
 PT compounds useful for treating mental, biological or medical diseases  
 XX

PS Disclosure; Page 29-30; 38pp; English.

XX The present sequence is a cDNA encoding novel human protein  
CC (NHP) known as human kinase. The human kinases share structural  
CC similarity with animal kinases, more particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.

XX Sequence 972 BP; 323 A; 172 C; 214 G; 263 T; 0 other;

Query Match 98.3%; Score 929; DB 22; Length 972;  
Best Local Similarity 100.0%; Pred. NO. 9.1e-241;  
Matches 929; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaagaatgataaaattagctgaagactggagaagggtctctatgggtgtattcaaa 60  
DB 1 atggaagaatgataaaattagctgaagactggagaagggtctctatgggtgtattcaaa 60  
QY 61 tcagagaaacaaacctctgcagaagttagctgttataaaattgtggaactcgaagat 120  
DB 61 tcagagaaacaaacctctgcagaagttagctgttataaaattgtggaactcgaagat 120  
QY 121 gactctgtgttaagaataatagcactaagaagaataacgtatgtgaagcaattaaacat 180  
DB 121 gactctgtgttaagaataatagcactaagaagaataacgtatgtgaagcaattaaacat 180  
QY 181 ccaaatctgtgaacctcctcagaggtgttcagagagaaaggaaatgcatttagttttt 240  
DB 181 ccaaatctgtgaacctcctcagaggtgttcagagagaaaggaaatgcatttagttttt 240  
QY 241 gaatactgtgatacatcacttttaaatgagctggaagaaacccaaatggattgctgat 300  
DB 241 gaatactgtgatacatcacttttaaatgagctggaagaaacccaaatggattgctgat 300  
QY 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttcgtcatalacat 360  
DB 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttcgtcatalacat 360  
QY 361 aactgtattcacagagatataaaacctgaaatattcttaataactaagcaaggaataac 420  
DB 361 aactgtattcacagagatataaaacctgaaatattcttaataactaagcaaggaataac 420  
QY 421 aagattgtgacttcgggtttgcaaaattctgattccaggagatgctaccaccgattat 480  
DB 421 aagattgtgacttcgggtttgcaaaattctgattccaggagatgctaccaccgattat 480  
QY 481 gtagctacagagatggtaccgagctcctgaactcttctgtggagatactcagtagttct 540  
DB 481 gtagctacagagatggtaccgagctcctgaactcttctgtggagatactcagtagttct 540  
QY 541 tcagtcgatataatgggtctattgtgtgttttttgcagagactcctgcagagccagcactg 600  
DB 541 tcagtcgatataatgggtctattgtgtgttttttgcagagactcctgcagagccagcactg 600

QY 601 tggcctggaaaatcagatgtggaccacactttatctgataataatcagaacactaggaataa 660  
DB 601 tggcctggaaaatcagatgtggaccacactttatctgataataatcagaacactaggaataa 660  
QY 661 atcccaagacatcaatcaatctttaaagtaacgggttttccatggcatcagtatacct 720  
DB 661 atcccaagacatcaatcaatctttaaagtaacgggttttccatggcatcagtatacct 720  
QY 721 gagccagaagacatggaacactcttgagaaaagtctcagatgttcatcctgtggtctcg 780  
DB 721 gagccagaagacatggaacactcttgagaaaagtctcagatgttcatcctgtggtctcg 780  
QY 781 aacttcataaggggtgtctgaagatgaatccagatgacagattaaacctgttcccaactc 840  
DB 781 aacttcataaggggtgtctgaagatgaatccagatgacagattaaacctgttcccaactc 840  
QY 841 ctggagagctcctactcttcttcttttcaagagcccaaatataaagaaagcagcgaat 900  
DB 841 ctggagagctcctactcttcttcttcttcaagagcccaaatataaagaaagcagcgaat 900  
QY 901 gaaggaagaacacagaagacgccaacaggt 929  
DB 901 gaaggaagaacacagaagacgccaacaggt 929  
RESULT 3  
AAS06725 ID AAS06725 standard; cDNA; 1083 BP.  
AC AAS06725;  
XX 12-SEP-2001 (first entry)  
DT Polynucleotide sequence encoding human protein kinase #25.  
DE Human; protein kinase; PK; STK; cancer; cardiovascular disease;  
KW metabolic disorder; immune related disease; neurological disorder;  
KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
KW reproductive disorder; gene therapy; ss.  
XX Homo sapiens.  
XX WO200138503-A2.  
XX 31-MAY-2001.  
XX 22-NOV-2000; 2000WO-US32085.  
XX 24-NOV-1999; 99US-0167482.  
XX (SUGE-) SUGEN INC.  
XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
PI Flanagan P, Clary D;  
XX WPT: 2001-343950/36.  
DR P-PSDB; AAU03525.  
XX Nucleic acids encoding human kinase polypeptides, useful for preventing  
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
PT neuronal-associated diseases, and microbial infections -  
XX Example 1; Figure 1; 433pp; English.  
XX AAS06701-AAS06757 encode for novel human protein kinases #1-57. The  
CC novel protein kinases have been identified as members of the tyrosine  
CC or serine/threonine kinase (PK and STK) families. The polynucleotides  
CC encoding protein kinases and the polypeptides may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate kinase expression. For example, they may be used to treat  
CC cancers (especially cancers of haematopoietic origin), cardiovascular  
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
CC immune related diseases (e.g. rheumatoid arthritis), neurological

CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
 CC Additionally, polynucleotides encoding protein kinases may be  
 CC used for gene therapy and as DNA probes in diagnostic assays.  
 CC The protein kinase polypeptides may be used as antigens in the production  
 CC of antibodies against the protein kinases and in assays to identify  
 CC modulators of protein kinase expression and activity.  
 XX  
 SQ Sequence 1083 BP; 366 A; 199 C; 225 G; 293 T; 0 other;

Query Match 91.3%; Score 863; DB 22; Length 1083;  
 Best Local Similarity 94.5%; Pred. No. 5.8e-223;  
 Matches 927; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

QY 1 atggaagaatgataaaatagctgaagctgagaagggcttattggtgtgtattccaaa 60  
 Db 1 atggaagaatgataaaatagctgaagctgagaagggcttattggtgtgtattccaaa 60

QY 61 tgcagaaacaaacctctgacaaagtagctgttaaaaaatttggaaatctgaagat 120  
 Db 61 tgcagaaacaaacctctgacaaagtagctgttaaaaaatttggaaatctgaagat 120

QY 121 gatcctgtgttaagaataagcactaaagaaatcgtatgttgaagcaattaaaacat 180  
 Db 121 gatcctgtgttaagaataagcactaaagaaatcgtatgttgaagcaattaaaacat 180

QY 181 ccaaatctgtgaacctatcgagggtgttcaggagaaagaaatgcattagtttt 240  
 Db 181 ccaaatctgtgaacctatcgagggtgttcaggagaaagaaatgcattagtttt 240

QY 241 gaatactgtatcacaacttttaattgaagctggaagaaacccaatgagttgtgat 300  
 Db 241 gaatactgtatcacaacttttaattgaagctggaagaaacccaatgagttgtgat 300

QY 301 ggagtgatcaaaagcgtattatgcaaacacttcaagctcttaattctgtcatatcat 360  
 Db 301 ggagtgatcaaaagcgtattatgcaaacacttcaagctcttaattctgtcatatcat 360

QY 361 aactgtattcacagagatataaaacctgaaatattcttaataacgaaggaataatc 420  
 Db 361 aactgtattcacagagatataaaacctgaaatattcttaataacgaaggaataatc 420

QY 421 aagattgtgacttcgggtgttcacaaattctga----- 454  
 Db 421 aagattgtgacttcgggtgttcacaaattctgagtaattcacaatgttgggagaattgac 480

QY 455 -----ttccaggagatgctacacccgattatgtacct 486  
 Db 481 ttacctgatctgatgtgactttgcagttccaggagatgctacacccgattatgtacct 540

QY 487 acgagatgttacagagctctgaaactcttggagagatactcagtatggtcttcagtc 546  
 Db 541 acgagatgttacagagctctgaaactcttggagagatactcagtatggtcttcagtc 600

QY 547 gatataatgggtctatgtgtgtttttgcagagctcctgacagccagccactgtggcct 606  
 Db 601 gatataatgggtctatgtgtgtttttgcagagctcctgacagccagccactgtggcct 660

QY 607 ggaataatcagatgtggaccaactttatctgataatcagaacactgagaaatattccca 666  
 Db 661 ggaataatcagatgtggaccaactttatctgataatcagaacactgagaaatattccca 720

QY 667 agacatcaatcaacttttaaaagtaacgggtttttccatggcatcagtataccctgagcca 726  
 Db 721 agacatcaatcaacttttaaaagtaacgggtttttccatggcatcagtataccctgagcca 780

QY 727 gaagacatgaaactcttgagaaagttctcagatgttcatcctgtggtcttgaaacttc 786  
 Db 781 gaagacatgaaactcttgagaaagttctcagatgttcatcctgtggtcttgaaacttc 840

QY 787 atgaaggggtgtctgaagatgaatccagatgacagattaacctgttcccaactcctggag 846

Db 841 atgaaggggtgtctgaagatgaatccagatgacagattaacctgttcccaactcctggag 900  
 QY 847 agctctactttgattcttttcaagggcccaatttaaaagaaagcacgtaataagaagga 906  
 Db 901 agctctactttgattcttttcaagggcccaatttaaaagaaagcacgtaataagaagga 960  
 QY 907 agaacaacagaaagacgccaacag 927  
 Db 961 agaacaacagaaagacgccaacag 981

RESULT 4  
 AAD03816  
 ID AAD03816 standard; cDNA; 1041 BP.  
 XX  
 AC AAD03816;  
 XX  
 DT 19-JUN-2001 (first entry)  
 XX  
 DE Human kinase cDNA #5.  
 XX  
 KW Human; kinase; gene therapy; bio reactor; mental disorder;  
 KW biological disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT I.1041 /\*tag= a  
 FT /\*product= "Human kinase #5"  
 FT /\*note= "The coding region does not include stop codon"  
 XX  
 PN WO200123579-A1.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 27-SEP-2000; 2000WO-US26621.  
 XX  
 PR 28-SEP-1999; 99US-0156511.  
 XX  
 PA (LEXI-) LEXICON GENETICS INC.  
 XX  
 PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
 XX  
 DR WPI: 2001-266166/27.  
 XX  
 P-PSDB; AA500494.  
 XX  
 PT New isolated human kinase polynucleotide useful for generating  
 PT antibodies, as reagents in diagnostic assays and for screening for  
 PT compounds useful for treating mental, biological or medical diseases -  
 XX  
 PS Claim 1; Page 31-32; 38pp; English.

CC The present sequence is a cDNA encoding novel human protein  
 CC (NHP) known as human kinase. The human kinases share structural  
 CC similarity with animal kinases, more particularly serine or  
 CC threonine protein kinases. Human kinase cDNA is useful for the  
 CC detection of mutant human kinase for the diagnosis of disease,  
 CC and also as a therapeutic. It is useful for screening drugs,  
 CC effective in the treatment of symptomatic or phenotypic  
 CC manifestations perturbing the normal function of NHP in the  
 CC body. The NHP nucleotide sequences are useful for generation of  
 CC antibodies, as reagents in diagnostic assays, for the  
 CC identification of other cellular gene products related to human  
 CC kinases, and as reagents in assays for screening compounds that  
 CC are useful for treating mental, biological or medical disorders.  
 CC NHP oligonucleotides are used as probes. The labelled NHP probes  
 CC are useful for screening human genomic library for identifying  
 CC polymorphisms and as primers in amplification assays to detect  
 CC mutations within the exons, introns and splice sites that can  
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct



CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
 CC encoding NHP products are used to genetically engineer cells  
 CC in vivo that functions as bioreactors in the body delivering a  
 CC continuous supply of NHP to the body. Nucleotide constructs  
 CC encoding functional NHPs are used in gene therapy for the  
 CC modulation of NHP expression.  
 XX  
 SQ Sequence 1068 BP; 344 A; 197 C; 240 G; 287 T; 0 other;

Query Match 87.1%; Score 823; DB 22; Length 1068;  
 Best Local Similarity 90.6%; Pred. No. 3.5e-212;  
 Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 1 atggaagaatgataaaattagctagctgagagctgagaggggtcttattggtgtgtattcaaa 60  
 Db 1 atggaagaatgataaaattagctagctgagaggggtcttattggtgtgtattcaaa 60  
 QY 61 tgcagaaacaaacctctgcacagtagtagctgttaaaaaattgtggaatctgaagat 120  
 Db 61 tgcagaaacaaacctctggacagtagtagctgttaaaaaattgtggaatctgaagat 120  
 QY 121 gatcctgttgtaagaaatagcactaagagaaatacgtatgttgaagcaataaaacat 180  
 Db 121 gatcctgttgtaagaaatagcactaagagaaatacgtatgttgaagcaataaaacat 180  
 QY 181 ccaaatctgtgaacctatcgaggtgttcagagaaaaagaaataacattagtttt 240  
 Db 181 ccaaatctgtgaacctatcgaggtgttcagagaaaaagaaataacattagtttt 240  
 QY 241 gaatactgtatcacaattttaaatgagctggaagaaacccaaatggagttgctgat 300  
 Db 241 gaatactgtatcacaattttaaatgagctggaagaaacccaaatggagttgctgat 300  
 QY 301 ggagtgatcaaaacgctattatggcaaaccttcaagctcttaatttctctatatacat 360  
 Db 301 ggagtgatcaaaacgctattatggcaaaccttcaagctcttaatttctctatatacat 360  
 QY 361 aactgtattcacagagatataaaacctgaaataattcttaataacgaaggaataatc 420  
 Db 361 aactgtattcacagagatataaaacctgaaataattcttaataacgaaggaataatc 420  
 QY 421 aagatttgacttcggtttgcaaaattctgtatccagagagatgcctacacattat 480  
 Db 421 aagatttgacttcggtttgcaaaattctgtatccagagagatgcctacacattat 480  
 QY 481 gtacgtacagatggtaccgagctcctgaactcttgggagatactcagttggtct 540  
 Db 481 gtacgtacagatggtaccgagctcctgaactcttgggagatactcagttggtct 540  
 QY 541 tcagtcgatatggctctattggtgtgttttgcagagctcctgacagccagccactg 600  
 Db 541 tcagtcgatatggctctattggtgtgttttgcagagctcctgacagccagccactg 600  
 QY 601 tggcctgaaatacagatgtggacaaacttctgataatcagaacact----- 650  
 Db 601 tggcctgaaatacagatgtggacaaacttctgataatcagaacact----- 650  
 QY 651 -----  
 Db 651 -----  
 QY 661 ggggttcgactgttgacaggctggtctgcgaactcttgagctcaagtgatccactgcc 720  
 Db 661 ggggttcgactgttgacaggctggtctgcgaactcttgagctcaagtgatccactgcc 720  
 QY 671 gtagctctcaagtgctggaattacagagaaataatcccaagacatacatcttt 684  
 Db 671 gtagctctcaagtgctggaattacagagaaataatcccaagacatacatcttt 684  
 QY 685 aagaataacgggttttccatgcacatcactgagccagaaagacatggaactctt 744  
 Db 685 aagaataacgggttttccatgcacatcactgagccagaaagacatggaactctt 744  
 QY 745 gaggaaaaagttcagatgttctcctgtggtctgaacttcatgaagggtgtctgaag 804  
 Db 745 gaggaaaaagttcagatgttctcctgtggtctgaacttcatgaagggtgtctgaag 804  
 QY 841 gaggaaaaagttcagatgttctcctgtggtctgaacttcatgaagggtgtctgaag 900  
 Db 841 gaggaaaaagttcagatgttctcctgtggtctgaacttcatgaagggtgtctgaag 900

QY 805 atgaatccagatgacagattaaactgttcccaactcctggagagctcctacttattct 864  
 Db 901 atgaatccagatgacagattaaactgttcccaactcctggagagctcctacttattct 960  
 QY 865 tttaagagggcccaattaaagaaaagcacgtatgaagaaagaaacagagacgcaa 924  
 Db 961 tttaagagggcccaattaaagaaaagcacgtatgaagaaagaaacagagacgcaa 1020  
 QY 925 caggt 929  
 Db 1021 caggt 1025  
 RESULT 6  
 AAF44669  
 ID AAF44669 standard; cDNA; 911 BP.  
 XX AC AAF44669;  
 XX XX  
 DT 27-MAR-2001 (first entry)  
 XX  
 XX Novel protein kinase cDNA, SEQ ID NO: 49.  
 DE Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
 KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;  
 KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;  
 KW immune disorder; cardiovascular disease; neurodegenerative disease;  
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.  
 XX Homo sapiens.  
 OS  
 XX WO200073469-A2.  
 PN  
 XX 07-DEC-2000.  
 PD  
 XX  
 PF 26-MAY-2000; 2000WO-US14842.  
 XX  
 PR 28-MAY-1999; 99US-0136503.  
 XX  
 XX (SUGEN) SUGEN INC.  
 PA  
 XX Plowman GB, Martinez R, Whyte D, Sudersanam S;  
 PI  
 XX WPI; 2001-032161/04.  
 DR  
 XX P-PSDB; AAB65642.  
 PT  
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
 treating immune-related diseases and disorders, cardiovascular disease,  
 neurodegenerative diseases and/or cancers -  
 PS  
 XX Example 4; Fig 2; 310pp; English.  
 CC  
 XX The present sequence encodes a novel protein kinase. The nucleic acids  
 and the protein kinases they encode may be used in the treatment and  
 diagnosis of diseases associated with inappropriate kinase expression  
 such as immune-related diseases and disorders, cardiovascular disease,  
 neurodegenerative diseases and/or cancers. The nucleic acids and  
 complementary sequences may also be used as DNA probes in diagnostic  
 assays. The kinase polypeptides may be used as antigens in the production  
 of antibodies of kinase expression and activity. Anti-kinase antibodies  
 and kinase antagonists may also be used to down regulate kinase  
 expression and activity. Diseases related to kinase expression and  
 activity include rheumatoid arthritis, atherosclerosis, autoimmune  
 disorders, complications of organ transplantation, myocardial infarction,  
 immune disorders, cardiomyopathies, strokes, renal failure,  
 oxidative-stress related disorders, chronic inflammatory bowel disease,  
 chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
 osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
 reproductive disorders.  
 CC  
 XX Sequence 911 BP; 293 A; 182 C; 178 G; 258 T; 0 other;  
 SQ



Db 418 ggcgcgtcgctgtgttttgcagagctcctgcagcgggtcagccactctggccgggaaatc 477  
 QY 615 agatgtggaccactttatctgataatcagaacactagaaaattaatcccaagacatca 674  
 Db 478 cgaagtggaccagctttacctgatcatcagagcgttgggaagctgattcccaagaccca 537  
 QY 675 atcaatctttaaagtaacgggtttttccatgcagtcagtcatactcctgagcagaagacat 734  
 Db 538 gctctctttagagtaaccagttttccgcgcgcagtcagtcatactcctgaaaccagaggacat 597  
 QY 735 ggaactcttgagaaagttctcagatgttctcctgtgctgtgctgaaactcagaaagg 794  
 Db 598 ggaactccttgagaaatctcgaattctcaattgttcagctgtgctttaaagtcttcagagg 657  
 QY 795 gctgtgaagatgaatccagatgacagattaaacctgttcccaactcctcggagactccta 854  
 Db 658 atgcctgaagatgaatcctgatgagggctgaacctgtgccagctgtcggacagtgctca 717  
 QY 855 ctttgattctttcaagagcccaattaaagaaagacscgtaatgaaggaagaacag 914  
 Db 718 ctttgagcttttcaagaggtacaatgaaagaaagccgcagtgaggggagagagccg 777  
 QY 915 aagacccaacagtgacttcgct 938  
 Db 778 aaggcgccagcagatcaactgct 801

## RESULT 8

AA03818  
 ID AA03818 standard; cDNA; 1819 BP.  
 XX  
 AC AA03818;  
 XX  
 DT 19-JUN-2001 (first entry)  
 XX  
 DE Human kinase cDNA #7.  
 XX  
 KW Human; kinase; gene therapy; bioreactor; mental disorder;  
 KW biological disorder; polymorphism; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT 5'UTR 1..137  
 FT /\*tag= a  
 FT CDS 138..701  
 FT /\*tag= b  
 FT /\*product= "Human kinase #1"  
 FT 3'UTR 702..1819  
 FT /\*tag= c  
 FT misc\_difference 1684  
 FT /\*tag= d  
 FT /\*note= "polymorphism site; given as r in the sequence"  
 XX  
 PN WO200123579-A1.  
 XX  
 XX 05-APR-2001.  
 PD  
 XX  
 XX 27-SEP-2000; 2000WO-US26621.  
 PF  
 XX  
 XX 28-SEP-1999; 99US-0156511.  
 PR  
 XX  
 PA (LEXI-) LEXICON GENETICS INC.  
 XX  
 PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
 XX  
 XX WPI; 2001-266166/27.  
 DR  
 DR P-PSDB; AAE00490.  
 XX  
 XX New isolated human kinase polynucleotide useful for generating  
 PT antibodies, as reagents in diagnostic assays and for screening for  
 PT compounds useful for treating mental, biological or medical diseases -

PS Disclosure; Page 34-35; 38pp; English.  
 XX  
 CC The present sequence encodes a novel human protein (NHP) which  
 CC is a human kinase. A polymorphism was identified in the 3' UTR  
 CC of the present sequence. The human kinases share structural  
 CC similarity with animal kinases, more particularly serine or  
 CC threonine protein kinases. Human kinase cDNA is useful for the  
 CC detection of mutant human kinase for the diagnosis of disease,  
 CC and also as a therapeutic. It is useful for screening drugs,  
 CC effective in the treatment of symptomatic or phenotypic  
 CC manifestations perturbing the normal function of NHP in the  
 CC body. The NHP nucleotide sequences are useful for generation of  
 CC antibodies, as reagents in diagnostic assays, for the  
 CC identification of other cellular gene products related to human  
 CC kinases, and as reagents in assays for screening compounds that  
 CC are useful for treating mental, biological or medical disorders.  
 CC NHP oligonucleotides are used as probes. The labelled NHP probes  
 CC are useful for screening human genomic library for identifying  
 CC polymorphisms and as primers in amplification assays to detect  
 CC mutations within the exons, introns and splice sites that can  
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
 CC encoding NHP products are used to genetically engineer cells  
 CC in vivo that functions as bioreactors in the body delivering a  
 CC continuous supply of NHP to the body. Nucleotide constructs  
 CC encoding functional NHPs are used in gene therapy for the  
 CC modulation of NHP expression.  
 XX  
 SQ Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;

Query Match 53.7%; Score 507; DB 22; Length 1819;  
 Best Local Similarity 69.8%; Pred. No. 6.5e-127;  
 Matches 929; Conservative 0; Mismatches 0; Indels 402; Gaps 2;  
 QY 1 atggaaaagtatgaaaattagctaaactggaaggggtcttatgggtgtgtatcaca 60  
 Db 138 atggaaaagtatgaaaattagctaaactggaaggggtcttatgggtgtgtatcaca 197  
 QY 61 tgcagaaaacaaacctctggcaagtagtgcgtgttaaaaatttgggaactcgaagat 120  
 Db 198 tgcagaaaacaaacctctggcaagtagtgcgtgttaaaaatttgggaactcgaagat 257  
 QY 121 gatcctgtgttaagaaaatagcactaagaaatagctatgttgaaagcaattaaacat 180  
 Db 258 gatcctgtgttaagaaaatagcactaagaaatagctatgttgaaagcaattaaacat 317  
 QY 181 ccaaatctgtgaacctcagcagggtgttcaggagaaaaggaaatgcatttagtttt 240  
 Db 318 ccaaatctgtgaacctcagcagggtgttcaggagaaaaggaaatgcatttagtttt 377  
 QY 241 gaatactgtatcatcacattttaaatgagctggaagaaagaaacccaaatggagtgctgat 300  
 Db 378 gaatactgtatcatcacattttaaatgagctggaagaaagaaacccaaatggagtgctgat 437  
 QY 301 ggagtgatcaaaaagcgtattatgcgaacacttcaagcttcttaattctgtcatatcat 360  
 Db 438 ggagtgatcaaaaagcgtattatgcgaacacttcaagcttcttaattctgtcatatcat 497  
 QY 361 aactgtattcacagagataaaaacctgaaaatattcttaactaaaggaagaaatc 420  
 Db 498 aactgtattcacagagataaaaacctgaaaatattcttaactaaaggaagaaatc 557  
 QY 421 agatttggacttcgggtttgcacaaattctg----- 453  
 Db 558 agatttggacttcgggtttgcacaaattctgagttggacttcttctctgtgtgccc 617  
 QY 454 ----- 453  
 Db 618 tcttgattgcttaattagttgaccttctgaattcttttctgccaattcagagatttt 677  
 QY 454 ----- 453  
 Db 678 ctctggttgatccattgctgacacagtgtttaccacatgggcccaggtcatctoga 737



QY 454 ----- 453  
Db 738 actttgacctcaagtgatccttcacactcgccctcccaaaagtctggaattgcaagtgtg 797  
QY 454 ----- 453  
Db 798 agccaccgtgccagccagatgttttcaacaataactactgagagctcacaagaattgttt 857  
QY 454 ----- attccaggagatgcctacacc 474  
Db 858 tiagtgggaacacaaatttcgaacaaattcttgagaacgcattccaggagatgcctacacc 917  
QY 475 gattatgtagctacagatggtaccagagctcctgaactctctgtgtggagatactcagttat 534  
Db 918 gattatgtagctacagatggtaccagagctcctgaactctctgtgtggagatactcagttat 977  
QY 535 gttcttcagtcgatatatgggctattggtgtgttttgcagagctcctgcagagccag 594  
Db 978 gttcttcagtcgatatatgggctattggtgtgttttgcagagctcctgcagagccag 1037  
QY 595 coactgtgacctggaataacagatgtggacacactttatctgaataatcagaacact---- 650  
Db 1038 ccaactgtgacctggaataacagatgtggacacactttatctgaataatcagaacactagta 1097  
QY 651 ----- 650  
Db 1098 gagacggggttccgcatgttgaccaggtgtctcgaaactcttgacgtcgaagtgtacca 1157  
QY 651 ----- agaaaaataatcccaagacatcaatca 678  
Db 1158 cctgcgtagctctcaaaagtgtggaattacaggagaaattatcccaagacatcaatca 1217  
QY 679 atctttaaagttaacgggttttccatggcatcagatatacctgagccagagacatggaa 738  
Db 1218 atctttaaagttaacgggttttccatggcatcagatatacctgagccagagacatggaa 1277  
QY 739 actcttgaggaagaaactctcagatgttcactctgtgctctgaactctatgaagggtgt 798  
Db 1278 actcttgaggaagaaactctcagatgttcactctgtgctctgaactctatgaagggtgt 1337  
QY 799 ctgaagatgaatccagatgacagatataactctgtcccaactcctgagagctcctacttt 858  
Db 1338 ctgaagatgaatccagatgacagatataactctgtcccaactcctgagagctcctacttt 1397  
QY 859 gatcttttcaagagcccaaatataaagaaagcacgtaatgaaggagaacagaaga 918  
Db 1398 gatcttttcaagagcccaaatataaagaaagcacgtaatgaaggagaacagaaga 1457  
QY 919 cgccaaacaggt 929  
Db 1458 cgccaaacaggt 1468

RESULT 9  
AAD03812 standard; cDNA, 561 BP.  
XX  
AC AAD03812;  
XX  
XX 19-JUN-2001 (first entry)  
DT  
XX Human kinase cDNA #1.  
DE  
XX Human; kinase: gene therapy; bioreactor; mental disorder;  
KW biological disorder; ss.  
KW  
XX Homo sapiens.  
OS  
XX Key  
FH 1..561  
FT CDS /\*tag= a  
FT /product= "Human kinase #1"

FT  
FT XX  
PN XX  
XX  
PD  
XX  
XX  
PF 27-SEP-2000; 2000WO-US26621.  
XX  
XX 28-SEP-1999; 99US-0156511.  
PR  
XX (LEXI-) LEXICON GENETICS INC.  
PA  
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
PI  
XX WPI: 2001-266166/27.  
DR  
XX P-PSDB; AAE00490.  
DR  
XX New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases -  
PS  
XX Disclosure: Page 27; 38pp; English.  
XX  
CC The present sequence is a cDNA encoding novel human protein  
CC (NHP) known as human kinase. The human kinases share structural  
CC similarity with animal kinases, more particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.  
XX  
SQ Sequence 561 BP; 189 A; 87 C; 112 G; 173 T; 0 other;

Query Match 48.1%; Score 454.4; DB 22; Length 561;  
Best Local Similarity 99.8%; Pred. No. 6.2e-113;  
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggaagaatgataaaaaattagtaagactgaagaggtcttatgggtgtattcaaa 60  
Db 1 atggaagaatgataaaaaattagtaagactgaagaggtcttatgggtgtattcaaa 60  
QY 61 tgcagaacacaaacctctggacaagttagctgttataaaatttgggaattcgaagt 120  
Db 61 tgcagaacacaaacctctggacaagttagctgttataaaatttgggaattcgaagt 120  
QY 121 gatcctgttttaagaaaaatgacctgaagaaatcgtatgttgaagcaattcaaacat 180  
Db 121 gatcctgttttaagaaaaatgacctgaagaaatcgtatgttgaagcaattcaaacat 180  
QY 181 ccaaatctgtgaacctcatcgaggtgttcaggagaaaaagaaagcatttagttttt 240  
Db 181 ccaaatctgtgaacctcatcgaggtgttcaggagaaaaagaaagcatttagttttt 240  
QY 241 gaatactgtatcatcaccttttaaatgagctggaagaaacccaaatggagttgctgat 300

Db 241 gaactactgtatcatacacttttaaatgagctggaagaacccaatggagtctgat 300  
 QY 301 ggagtgataaagcgattatggcaaacacttcaagctcttaattctgtcatatacat 360  
 Db 301 ggagtgataaagcgattatggcaaacacttcaagctcttaattctgtcatatacat 360  
 QY 361 aactgtattcacagagatataaaacctgaaatattcttaataactaagcaaggaaatac 420  
 Db 361 aactgtattcacagagatataaaacctgaaatattcttaataactaagcaaggaaatac 420  
 QY 421 aagatttggacttcgggttggcacaattctgatt 456  
 Db 421 aagatttggacttcgggttggcacaattctgatt 456

## RESULT 10

AAD03815

ID AAD03815 standard; cDNA; 594 BP.

XX AC AAD03815;

XX DT 19-JUN-2001 (first entry)

XX DE Human kinase cDNA #4.

XX KW Human; kinase; gene therapy; bioreactor; mental disorder;

XX KW biological disorder; ss.

XX OS Homo sapiens.

XX FH Key

XX FT COS

XX FT Location/Qualifiers

XX FT 1..594

XX FT /\*tag= a

XX FT /product= "Human kinase #4"

XX FT /note= "The coding region does not include stop codon"

XX FT /partial

XX PN W0200123579-Al.

XX PD 05-APR-2001.

XX PF 27-SEP-2000; 2000WO-US26621.

XX PR 28-SEP-1999; 99US-0156511.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;

XX DR WPI: 2001-266166/27.

XX DR P-PSDB; AAE00493.

XX PT New isolated human kinase polynucleotide useful for generating

XX PT antibodies, as reagents in diagnostic assays and for screening for

XX PT compounds useful for treating mental, biological or medical diseases -

XX PS Disclosure; Page 31; 38pp; English.

XX CC The present sequence is a cDNA encoding novel human protein

XX CC (NHP) known as human kinase. The human kinases share structural

XX CC similarity with animal kinases, more particularly serine or

XX CC threonine protein kinases. Human kinase cDNA is useful for the

XX CC detection of mutant human kinase for the diagnosis of disease,

XX CC and also as a therapeutic. It is useful for screening drugs,

XX CC effective in the treatment of symptomatic or phenotypic

XX CC manifestations perturbing the normal function of NHP in the

XX CC body. The NHP nucleotide sequences are useful for generation of

XX CC antibodies, as reagents in diagnostic assays, for the

XX CC identification of other cellular gene products related to human

XX CC kinases, and as reagents in assays for screening compounds that

XX CC are useful for treating mental, biological or medical disorders.

XX CC NHP oligonucleotides are used as probes. The labelled NHP probes

XX CC are useful for screening human genomic library for identifying

CC polymorphisms and as primers in amplification assays to detect  
 CC mutations within the exons, introns and splice sites that can  
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
 CC encoding NHP products are used to genetically engineer cells  
 CC in vivo that functions as bioreactors in the body delivering a  
 CC continuous supply of NHP to the body. Nucleotide constructs  
 CC encoding functional NHPs are used in gene therapy for the  
 CC modulation of NHP expression.  
 XX

SQ Sequence 594 BP; 203 A; 94 C; 115 G; 182 T; 0 other;

## Query Match

Best Local Similarity 48.1%; Score 454.4; DB 22; Length 594;

Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggaaaagtatgaaaattagctaaagactggagaagggtctctatgggtgtattcaaa 60

Db 1 atggaaaagtatgaaaattagctaaagactggagaagggtctctatgggtgtattcaaa 60

QY 61 tgcagaaacaaaacctctctggacaagtagctgtttaaaaaatttgggaatctgaagat 120

Db 61 tgcagaaacaaaacctctctggacaagtagctgtttaaaaaatttgggaatctgaagat 120

QY 121 gatcctgtgttaagaaaatagcactaagaagaatacgtatgttgaagcaattaaaaaat 180

Db 121 gatcctgtgttaagaaaatagcactaagaagaatacgtatgttgaagcaattaaaaaat 180

QY 181 ccaaatcttgtgaacctatcgaggtgttcaggagaaaagaaatgcatttagtttt 240

Db 181 ccaaatcttgtgaacctatcgaggtgttcaggagaaaagaaatgcatttagtttt 240

QY 241 gaatactgtgatcatcacactttttaaagtgcgtgaaagaaacccaatgagtgctgat 300

Db 241 gaatactgtgatcatcacactttttaaagtgcgtgaaagaaacccaatgagtgctgat 300

QY 301 ggagtgataaagcgattatggcaaacacttcaagctcttaattctgtcatatacat 360

Db 301 ggagtgataaagcgattatggcaaacacttcaagctcttaattctgtcatatacat 360

QY 361 aactgtattcacagagatataaaacctgaaatattcttaataactaagcaaggaaatac 420

Db 361 aactgtattcacagagatataaaacctgaaatattcttaataactaagcaaggaaatac 420

QY 421 aagatttggacttcgggttggcacaattctgatt 456

Db 421 aagatttggacttcgggttggcacaattctgatt 456

## RESULT 11

AAA29746

ID AAA29746 standard; DNA; 1698 BP.

XX AC AAA29746;

XX DT 15-AUG-2000 (first entry)

XX DE Rabbit KKIAMRE kinase genomic DNA SEQ ID NO:5.

XX KW Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;

XX KW cdc2-related kinase; brain; gene therapy; genetic disorder; detection;

XX KW identification; ds.

XX OS Oryctolagus cuniculus.

XX PN W0200020567-A2.

XX PD 13-APR-2000.

XX PF 01-OCT-1999; 99WO-US23010.

XX PR 02-OCT-1998; 98US-0102906.

XX

PA (UYSC-) UNIV SOUTHERN CALIFORNIA.  
XX Thompson RF, Gomi H, Sun W;  
PI WPI; 2000-328932/28.  
XX Novel learning induced kinase polynucleotides and polypeptides, useful  
PT for the analysis of learning and memory, and for gene therapy -  
XX Claim 1; Fig 9; 64pp; English.  
XX The present sequence encodes a learning-induced kinase, designated  
CC KKIAMRE kinase, which is isolated from rabbit brain tissue. KKIAMRE  
CC kinase is a cdc2-related kinase. The KKIAMRE kinase polynucleotides can  
CC be used to express recombinant protein for analysis, characterisation or  
CC therapeutic use, as markers for tissues in which the protein is  
CC preferentially expressed, as molecular weight markers on Southern gels,  
CC as chromosome markers or tags, to compare endogenous DNA sequences in  
CC patients to identify potential genetic disorders, as probes to hybridise  
CC an antigen to induce anti-DNA antibodies. The polypeptides can be used  
CC in assay to discover biological activity, to raise antibodies, and as  
CC markers, and to isolate correlative receptors or ligands. The  
CC polynucleotides may also be used for gene therapy for the treatment of  
CC disorders which are mediated by KKIAMRE kinase.  
XX Sequence 1698 BP; 551 A; 311 C; 390 G; 446 T; 0 other;

Query Match 30.4%; Score 287; DB 21; Length 1698;  
Best Local Similarity 59.1%; Pred. No. 1.3e-67;  
Matches 510; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

Qy 1 atcgaaatgataaaattagctgaagactgagagggtctcttatgggtgtgtatcca 60  
Db 1 atcgaaatgataaaattagctgaagactgagagggtctcttatgggtgtgtatcca 60

Qy 61 tgcagaaacaaaccttgagacagtagctgtttaaataattgtgactcgaagat 120  
Db 61 tgcagaaacaaaccttgagacagtagctgtttaaataattgtgactcgaagat 120

Qy 121 gacccgtgtgttaagaaatagcactaagaaatcagttatgttgagcaatataa 180  
Db 121 gacccgtgtgttaagaaatagcactaagaaatcagttatgttgagcaatataa 180

Qy 181 ccaaatctgtgaacctcatcgaggtgttcaggagaaagaaagaaatgattttt 240  
Db 181 ccaaatctgtgaacctcatcgaggtgttcaggagaaagaaagaaatgattttt 240

Qy 241 gaatactgtgatacacttttaaatgagctggaagaaacccaatggagtgtgat 300  
Db 241 gaatactgtgatacacttttaaatgagctggaagaaacccaatggagtgtgat 300

Qy 301 ggaagtgcataaagctattatggcaaacactcgaagctcttaatttctgtcata 360  
Db 301 ggaagtgcataaagctattatggcaaacactcgaagctcttaatttctgtcata 360

Qy 361 aactgtattccagagatataaaacctgaaataattcttaataaagcaagaaata 420  
Db 361 aactgtattccagagatataaaacctgaaataattcttaataaagcaagaaata 420

Qy 421 aagattgtgacttcgggtttgcacaaattctg---attccagagagctgctacac 477  
Db 421 aagattgtgacttcgggtttgcacaaattctg---attccagagagctgctacac 477

Qy 478 tatgtagctacagatggtaccagagctcctgaaactcttgggtggagatactcag 537  
Db 478 tatgtagctacagatggtaccagagctcctgaaactcttgggtggagatactcag 537

Qy 538 tcttcagtcgatatatgggtctattgtgtgttttttgcagagctcctgacaggc 597  
Db 538 tcttcagtcgatatatgggtctattgtgtgttttttgcagagctcctgacaggc 597

Qy 541 aaagctgtgagatgtggtggccattgtgtgtctgtgtaactgaaatgctcatgc 600  
Db 541 aaagctgtgagatgtggtggccattgtgtgtctgtgtaactgaaatgctcatgc 600

Qy 598 ctgtgctgggaaaaatcagatgtggaccactttatctatgataatcaaacactagaaa 657  
Db 601 ctgttctctggagactgtatattgacagcttattctattatctatgaggtgttaagta 660

Qy 658 ttaatcccaagacatcaatcaatctttaaagtaacgggtttttcccatgcatcag 717  
Db 661 ctaattccaagacacacagagctttttataaaaatcctgtgttgcgagtaagg 720

Qy 718 cctgagccagaagacatggaactcttgagaaaagtctcagatgttcatcctgtg 777  
Db 721 cctgaaatcaaggaatcagaacctcttgaaagacgctatcccaagctctcaga 780

Qy 778 ctgaacttcataaggggtgtctgaagatgaatccagatgacagattaaacctgttcc 837  
Db 781 atagatttagcaagaagatgcttacatgttgaccagacaaaaggccctctgtg 840

Qy 838 ctctcggagagctcctactttga 860  
Db 841 ctcttaaccatgattctttca 863

RESULT 12  
AAA29745  
ID AAA29745 standard; cDNA; 3080 BP.  
XX AAA29745;  
XX  
XX 15-AUG-2000 (first entry)  
XX Rabbit KKIAMRE kinase encoding cDNA SEQ ID NO:3.  
XX Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;  
KW cdc2-related kinase; brain; gene therapy; genetic disorder; detection;  
KW identification; ss.  
XX Oryctolagus cuniculus.  
XX Key Location/Qualifiers  
FT CDS 249..1949  
FT /\*Tag= a  
FT /product= "KKIAMRE kinase"  
XX  
XX WO200020567-A2.  
XX  
XX 13-APR-2000.  
XX  
XX 01-OCT-1999; 99WO-US23010.  
XX  
XX 02-OCT-1998; 98US-0102906.  
XX (UYSC-) UNIV SOUTHERN CALIFORNIA.  
XX Thompson RF, Gomi H, Sun W;  
XX WPI; 2000-328932/28.  
XX P-PSDB; AAY90724.  
XX Novel learning induced kinase polynucleotides and polypeptides, useful  
PT for the analysis of learning and memory, and for gene therapy -  
XX Claim 1; Fig 4; 64pp; English.  
XX The present sequence encodes a learning-induced kinase, designated  
CC KKIAMRE kinase, which is isolated from rabbit brain tissue. KKIAMRE  
CC kinase is a cdc2-related kinase. The KKIAMRE kinase polynucleotides can  
CC be used to express recombinant protein for analysis, characterisation or  
CC therapeutic use, as markers for tissues in which the protein is  
CC preferentially expressed, as molecular weight markers on Southern gels,  
CC as chromosome markers or tags, to compare endogenous DNA sequences in  
CC patients to identify potential genetic disorders, as probes to hybridise  
CC an antigen to induce anti-DNA antibodies. The polypeptides can be used  
CC in assay to discover biological activity, to raise antibodies, and as  
CC markers, and to isolate correlative receptors or ligands. The  
CC polynucleotides may also be used for gene therapy for the treatment of  
CC disorders which are mediated by KKIAMRE kinase.

CC markers, and to isolate correlative receptors or ligands. The  
CC polynucleotides may also be used for gene therapy for the treatment of  
CC disorders which are mediated by KKLAMRE kinase.  
XX  
SQ Sequence 3080 BP; 906 A; 569 C; 765 G; 840 T; 0 other;

```
Query Match      30.4%; Score 287; DB 21; Length 3080;
Best Local Similarity 59.1%; Pred. No. 1.6e-67;
Matches 510; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

QY 1 atggaagaatgataaaattagctaaagactgagagagagcttcttcttggtgtgtattcaaa 60
   ||||| ||||| || || || || || || || || || || || || || || || || ||
Db 249 atggaagaatgagaaacctggtggtggtggtggtggtggtggtggtggtggtggtggt 308
   || || || || || || || || || || || || || || || || || || || || ||

QY 61 tgcagaaacaaacctctgacagtagtagctgtttaaataatttggaaatctgagat 120
   || || || || || || || || || || || || || || || || || || || || ||
Db 309 tglaggaataaagatagtggagaatttggccatcaagaagtctctagaagtgatgat 368
   || || || || || || || || || || || || || || || || || || || || ||

QY 121 gatcctgtgttaagaataatgacactaagagaaatagctatgttgaagcaattaaacat 180
   || || || || || || || || || || || || || || || || || || || || ||
Db 369 gacaaaatgtttaaaaaattgctatgcgagaatcaagttaactaaagcaactgagat 428
   || || || || || || || || || || || || || || || || || || || || ||

QY 181 ccaaatctgtgaacctcatcaggtgttcaggagaaaaaggaataatgcatttagtttt 240
   ||||| ||||| || || || || || || || || || || || || || || || || ||
Db 429 gaataattgttggaatctgttggtggtgtgtaaaaaaaacacggtgacactgacttt 488
   || || || || || || || || || || || || || || || || || || || || ||

QY 241 gaactctgtatcacatcattttaaatgagctggaagaaacccaaatggaggtgtgat 300
   || || || || || || || || || || || || || || || || || || || || ||
Db 489 gaattgttgaccacagactctctgatgaattggaaactcttccaaatgactagatgac 548
   || || || || || || || || || || || || || || || || || || || || ||

QY 301 ggaagtatcaaaaacgtattatggaacacctcaagctcttaattctgtcatatcat 360
   || || || || || || || || || || || || || || || || || || || || ||
Db 549 caagtatgctaaagattgttttcagattattaatggaattggatttgcacagtcac 608
   || || || || || || || || || || || || || || || || || || || || ||

QY 361 aactgtattccagagatataaaacctgaaataattcttaataactaagcaagataatc 420
   || || || || || || || || || || || || || || || || || || || || ||
Db 609 aatatcatacatagagatataaagccagagaatatatgtgtccagctgtcgtgtgc 668
   || || || || || || || || || || || || || || || || || || || || ||

QY 421 aagattgtgaactcgggttgacaaatctcg---attccagagagactcctacaccgat 477
   || || || || || || || || || || || || || || || || || || || || ||
Db 669 aagttatgtatttggatttgcagcgacactggcagctcccgagaggtttacactgat 728
   || || || || || || || || || || || || || || || || || || || || ||

QY 478 tatgtactacagatgtgtaccagactcctgaactcttctgtggagagatactcagtagt 537
   ||||| ||||| || || || || || || || || || || || || || || || || ||
Db 729 tatgtggcaactcagtggtacagagctccagaaactactggtgtgtgtatgcgaagtatg 788
   || || || || || || || || || || || || || || || || || || || || ||

QY 538 tcttcagtcgatatatgggctatgtgtgtgttttgcagagctcctgcagagccagcca 597
   || || || || || || || || || || || || || || || || || || || || ||
Db 789 aaagctgtgatgtgtggtggttctgtctgtgtaactgaaatgctcatggtggggaacc 848
   || || || || || || || || || || || || || || || || || || || || ||

QY 598 ctgtggtcctggaaaatcagatgtggaccacttctatctgataatcagaacactagaaaa 657
   ||||| ||||| || || || || || || || || || || || || || || || || ||
Db 849 ctgttctcctggagactctgatattgatcagctttatcttattatgaggtgttaagtaat 908
   || || || || || || || || || || || || || || || || || || || || ||

QY 658 ttaatcccaagacatacaatcaatctttaaagtaacgggtttttccatggcatcagata 717
   ||||| ||||| || || || || || || || || || || || || || || || || ||
Db 909 cttaattccaagacacagagagctttttataaaatcctctgtgttctggagtaaggttg 968
   || || || || || || || || || || || || || || || || || || || || ||

QY 718 cctgagcagaagaatggaactctcttgaggaagaattctcagatgttccatcctgtggt 777
   ||||| ||||| || || || || || || || || || || || || || || || || ||
Db 969 cctgaaatcaaggaatcagaacctcttgaaagacgctatcccgaagctctcagaagtgtg 1028
   || || || || || || || || || || || || || || || || || || || || ||

QY 778 ctgaactctcatgaaggggtgtctgagagatgaatccagatgacagattaacctgtcccaa 837
   || || || || || || || || || || || || || || || || || || || || ||
Db 1029 atagatttagcaagaataatgcttaccatgttgaccagacaaaaagccctctgtgctgag 1088
   || || || || || || || || || || || || || || || || || || || || ||

QY 838 ctctcggagcctcactcttga 860
   ||||| ||||| || || || || || || || || || || || || || || || || ||
Db 1089 ctctcacaccatgatttcttca 1111
   ||||| ||||| || || || || || || || || || || || || || || || || ||
```

RESULT 13

```
AAZ51208
ID AAZ51208 standard; cDNA; 1513 BP.
XX
AC AAZ51208;
XX
DI 06-JUN-2000 (first entry)
XX
DE Human lost in leukaemia kinase (LLK) cDNA.
XX
KW Lost in leukaemia kinase; LLK; cdc2-related kinase; human; leukaemia;
KW MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;
KW chromosome 5q31.1; cytostatic; mitotic index; treatment; prevention;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1368
FT /tag= a
FT /product= "Human lost in leukaemia kinase protein"
PN WO200012719-A1.
XX
PD 09-MAR-2000.
XX
PF 31-AUG-1999; 99WO-CA00794.
XX
PR 31-AUG-1998; 98CA-2243784.
PR 20-NOV-1998; 98CA-2251249.
XX
PA (ONTA-) ONTARIO CANCER INST.
XX
PI Zanke B, Haq R, Randall S, Midmer M;
XX
DR WPI; 2000-237880/20.
XX
P-PSDB; AAY70126.
XX
PT Isolated polynucleotide encoding a lost in leukemia kinase (LLK)
PS protein, useful for treatment, diagnosis and prevention of leukemia -
XX Claim 4c; Page 51; 69pp; English.
XX
CC The present sequence is the cDNA encoding human lost in leukaemia
CC kinase (LLK), isolated from human lambda gt10 foetal heart genomic
CC library. This gene is mapped to chromosome 5q31.1, a region implicated
CC in human acute leukaemia. It is expressed strongly in muscle, heart,
CC liver, brain and in tissues which have a very low mitotic index. LLK
CC has 83% amino acid sequence homology to rat LLK beta protein. LLK
CC is closely related to cdc2-related kinases, that are putative tumour
CC suppressors and to mitogen activated protein kinases (MAPKs). This
CC sequence has cytostatic activity. It is useful for the treatment,
CC diagnosis and prevention of acute leukaemia and is also used in gene
CC therapy.
XX
SQ Sequence 1513 BP; 551 A; 253 C; 295 G; 414 T; 0 other;

Query Match      24.1%; Score 227.4; DB 21; Length 1513;
Best Local Similarity 55.6%; Pred. No. 1.4e-51;
Matches 479; Conservative 0; Mismatches 376; Indels 6; Gaps 2;

QY 1 atggaagaatgataaaattagctaaagactgagagagcttcttcttggtgtgtattcaaa 60
   ||||| ||||| || || || || || || || || || || || || || || || || ||
Db 1 atggaagtgtatgaacctctgaaaccttggaagtggagagggaagtacggacagtcagaaa 60
   || || || || || || || || || || || || || || || || || || || || ||

QY 61 tgcagaaacaaacctctgacaaatagtagctgtttaaataatttggaaatctgagat 120
   || || || || || || || || || || || || || || || || || || || || ||
Db 61 tgtaaacataagaatactcggcagatagtaggccaattaagatatatttatgagagaccagaa 120
   || || || || || || || || || || || || || || || || || || || || ||

QY 121 gatcctgtgttgaagaaaaatagcactaagagaaatcagtagtctggaagcaattaaacat 180
   || || || || || || || || || || || || || || || || || || || || ||
Db 121 caatctgtc---aacaataatgcgatgagagaaaaaagtttcttaagaacatttctaac 177
   || || || || || || || || || || || || || || || || || || || || ||
```



Db 478 tatgtgccaacgctgggtatagagctcccgaaattagattataaaagatacttcttatgga 537  
QY 538 tcttcagtcgatataatgggtgattggtgttttggcagagctcctgacagccagcca 597  
Db 538 aaacctggtgatctcgtggttggctgtatgacatgagatggcactggaaatccc 597  
QY 598 ctgtgacctggaataatcagatgtaggacaaactttatctgataatcgaacactaggaaa 657  
Db 598 tatcttcctagtagtctgattggtatttactccataaaattgtttgaaagtggcaat 657  
QY 658 ttaatcccaagacatacaatacttttaaaagtaacgggtttttccatgcatcagata 717  
Db 658 ttgtccactccacttgagaatacttttccaagagccccaatttttgcgtgggtagtctt 717  
QY 718 cctgagccagaagacatgaaactcttgaggaaaagtctcagatgttcatcctgtgct 777  
Db 718 cctcaagtccaaccccccaaaaatcgaagaadaaataatccaagcttaattggattgtg 777  
QY 778 ctgaacttcataaggggtgtctgaagatgaatccagatgacagattaaacctgttccaa 837  
Db 778 gcagatatagttcatgctgtttacaaaattgatctctgctgacaggaatcatcatagt 837  
QY 838 ctctggagagctcctacttt 858  
Db 838 cttttgcatcatgagtatttt 858

RESULT 15

AAZ51207  
ID AAZ51207 standard; cDNA; 1667 BP.  
XX  
AC AAZ51207;  
XX  
DT 06-JUN-2000 (first entry)  
XX  
DE Rat lost in leukaemia kinase beta isoform (LLK beta) cDNA.  
XX  
KW Lost in leukaemia kinase; LLK beta; rat; cdc2-related kinase; leukaemia;  
KW MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;  
KW cytostatic; mitotic index; treatment; prevention; gene therapy; ss.  
XX  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT 64..1437  
FT CDS  
FT /\*tag= a  
FT /product= "Rat lost in leukaemia kinase beta protein"  
PN W0200012719-A1.  
XX  
XX  
PD 09-MAR-2000.  
XX  
PF 31-AUG-1999; 99WO-CA00794.  
XX  
PR 31-AUG-1998; 98CA-2243784.  
PR 20-NOV-1998; 98CA-2251249.  
XX  
PA (ONTA-) ONTARIO CANCER INST.  
XX  
PI Zanke B, Haq R, Randall S, Midmer M;  
XX  
DR WPI; 2000-237880/20.  
DR P-PSDB; AAY70125.  
XX  
XX Isolated polynucleotide encoding a lost in leukemia kinase (LLK)  
PT protein, useful for treatment, diagnosis and prevention of leukemia -  
XX  
PS Claim 4b; Page 48; 69pp; English.  
XX  
CC The present sequence is the cDNA encoding the beta isoform of rat lost in  
CC leukemia kinase (LLK beta), isolated from rat jejunum cDNA library. It  
CC is expressed strongly in muscle, heart, liver, brain and in tissues which

CC have a very low mitotic index. Rat LLK beta protein is closely related to  
CC cdc2-related kinases, that are putative tumour suppressors and to mitogen  
CC activated protein kinases (MAPKs). This sequence has cytoskeletal activity.  
CC It is useful for the treatment, diagnosis and prevention of acute  
CC leukaemia and is also used in gene therapy.  
XX

SQ Sequence 1667 BP; 525 A; 353 C; 392 G; 397 T; 0 other;

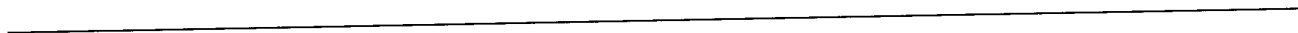
Query Match 22.4%; Score 211.4; DB 21; Length 1667;  
Best Local Similarity 54.5%; Pred. No. 3e-47;  
Matches 469; Conservative 0; Mismatches 386; Indels 6; Gaps 2;

QY 1 atggaataatgataaaattagctaaagactggaagaggtcttatgggtgtattcaaa 60  
Db 64 atggagatgataaaaccttgaaaaagtagggaggggaagtattggaacggtcatgaaa 123  
QY 61 tgcagaacaacaaacctctggacaagtagctgtttaaataattgtggaactctgaagat 120  
Db 124 tgcagcatgaaggtatctggcggtataggccattaaagtattctatgagaacccagaa 183  
QY 121 gatcctgtgttgaagaaatagcactaagagaaatacgtatgttgaacaaataaaacat 180  
Db 184 aaatctgtc---aacaaaattgcaacgagagaaataaaagtcttcaagcaattctcgt 240  
QY 181 ccaaatctgtgaacctcctcaggtgttcaggagaaaaagaaaatcttagttttt 240  
Db 241 gaaaacctgtcactctgattgaaagtttttagacaaaaaagaaaatccatttggattt 300  
QY 241 gaactgtgtatcacacttttaaatgagctggaagaaaccccaaatgagttgctgagt 300  
Db 301 gagttattgaccacacggtcttagatgagttacagcattattgtcacgattagagagt 360  
QY 301 ggagtgatacaaaagcgtattatgcaaacacttcaagctcttaattcttgcataacat 360  
Db 361 aagcagtaagaagaactctgtccagatctccgagcgtcagtagtatctgcacacaaac 420  
QY 361 aactgtattcacagagatataaaacctgaaataattcttaatacgaagaagaataac 420  
Db 421 aacattatcacccagatataaagcctgagaatttttagtctccagtcgaagaattaca 480  
QY 421 aagatttgaactcgggtttgcacaaattct---gattccagagaagctcctacacagat 477  
Db 481 aagctctggattttgggttttggcgagacactagcagctcctggagagcgtttcacagac 540  
QY 478 tatgtacacagatgataccgagctcctgaactcttctgtgggagatactcagtatggt 537  
Db 541 taagtggccacccctggcagagctccagagctggtgttgaaagacaccacctatgga 600  
QY 538 tcttcagtcgatataatgggtattggtgttttggcagagctcctgacagccagcca 597  
Db 601 aaaccagtgatctcgtgttgggtctgtatgacattgaaatggcactggcaatccc 660  
QY 598 ctgtggcctggaaaatcagatgtagaacacttttatctgataatcgaacactaggaaaa 657  
Db 661 taacttcctagcaggttccgatttggatttgcctccacaagattgttttaaaagtaggcaac 720  
QY 658 ttaatcccaagacataatcaatctttaaagaatacgggtttttccatggcatcagata 717  
Db 721 ctgaccccgacctgcacaataatttttccaagagtcccatcttctgctgggtgttctt 780  
QY 718 cctgagccagagacatggaaactcttgaggaaaagttctcagatgttcatcctgtgct 777  
Db 781 cctcaagtccaacatcccaaaaacgaagaagaataaccacaagctcaacggtatgctg 840  
QY 778 ctgaacttcataaggggtgtctgaaagtgaatccagatgacagattaacctgttcccaa 837  
Db 841 gcagatatagttcatgctgtgtttacaaaattgatctcgtgagaggaataatccaccgat 900  
QY 838 ctctggagagctcctacttt 858  
Db 901 cttttgcatcatgagtatttt 921

us-09-671-050-11.rng

Fri May 3 11:02:07 2002

Search completed: May 2, 2002, 17:13:19  
Job time: 5456 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:06:37 ; Search time 2235.62 Seconds  
(without alignments)  
4542.257 Million cell updates/sec

Title: US-09-671-050-11  
Perfect score: 945  
Sequence: 1 atggaaagtatgaaatt.....aggctcttcgcctcaaaagt 945

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_htc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_htc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rpd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	401.8	42.5	1691	12 AK016781	AK016781 Mus muscu
2	333.6	35.3	500	10 AI385966	AI385966 ml34h09.y
3	316.8	33.5	938	10 AA061797	AA061797 ml34h09.r
4	258.6	27.4	731	10 AW106692	AW106692 um32d03.y
5	255	27.0	718	10 AA286088	AA286088 vc33b05.r
6	209.8	22.2	579	10 AW233105	AW233105 fj28b10.y
7	200.4	21.2	295	10 AA626859	AA626859 zu89f10.s
8	192.4	20.4	1535	12 BC010966	BC010966 Homo sapi
9	186.6	19.7	727	10 AI508835	AI508835 vc33b05.y
10	174.8	18.5	582	10 AV672492	AV672492 AV672492
11	174.4	18.5	665	13 AQ953706	AQ953706 Sheared D
12	170.6	18.1	399	10 BE700387	BE700387 QV3-NN009

13	165.6	17.5	401	10 BE700389	BE700389 QV3-NN009
14	158.4	16.8	495	10 AI385990	AI385990 ml50h06.y
15	155.4	16.4	384	10 BE223187	BE223187 kp73q08.y
16	154.6	16.4	744	11 BF787339	BF787339 602113731
17	154	16.3	879	11 BG281841	BG281841 602403067
18	154	16.3	926	10 AL524036	AL524036 AL524036
19	150.4	15.9	779	11 BG760621	BG760621 602717373
20	146	15.4	1112	10 AT002436	AT002436 AT002436
21	145.4	15.4	628	10 AW611190	AW611190 un78e06.y
22	142.8	15.1	808	10 BI476896	BI476896 ul08g10.y
23	141.4	15.0	957	11 BI410031	BI410031 602962512
24	141.2	14.9	795	10 BE275968	BE275968 601120739
25	136.2	14.4	754	11 BG757858	BG757858 602711380
26	135.2	14.3	699	10 AL584951	AL584951 AL584951
27	135.2	14.3	752	10 BE746563	BE746563 601590135
28	135.2	14.3	911	11 BG397188	BG397188 602434149
29	134.8	14.3	1131	10 BE563495	BE563495 601334976
30	134.2	14.2	764	11 BG162489	BG162489 df15e08.y
31	134	14.2	202	10 BE135229	BE135229 ug28e10.y
32	134	14.2	696	11 BF479244	BF479244 L48-2906T
33	133	14.1	680	10 AL585015	AL585015 AL585015
34	132.8	14.1	619	11 BE844514	BE844514 EST259 AP
35	131.8	13.9	1000	11 BF316057	BF316057 601896036
36	131.4	13.9	831	11 BG825471	BG825471 602749632
37	131.2	13.9	884	11 BG440417	GA_Ea000
38	131	13.9	297	10 AI965145	AI965145 fc87e07.y
39	131	13.9	613	10 AL504123	AL504123 AL504123
40	131	13.9	778	10 BE413170	BE413170 MCG014.CO
41	131	13.9	963	11 BG420755	BG420755 602448546
42	130.6	13.8	732	10 AW701949	AW701949 ug93b12.y
43	130.2	13.8	793	11 BG700342	BG700342 603680030
44	130.2	13.8	796	11 BG915878	BG915878 602813329
45	130	13.8	713	10 AI246661	AI246661 ql27c01.x

ALIGNMENTS

RESULT 1

AK016781	1691 bp	mRNA	05-JUL-2001
LOCUS	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933411017, full insert sequence.		
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933411017, full insert sequence.		
ACCESSION	AK016781.1 GI:12855701		
VERSION	2		
KEYWORDS	CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 1691)		
TITLE	Carninci P. and Hayashizaki Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Methods in enzymology. 303, 19-44 (1999)		
PUBMED	95279253		
REFERENCE	2 (bases 1 to 1691)		
AUTHORS	Carninci P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome research. 10 (10), 1617-1630 (2000)		
MEDLINE	20493374		
PUBMED	11042159		
REFERENCE	3 (bases 1 to 1691)		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,		



B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter  
, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
Waterson, R. and Wilson, R.  
The WashU-NCI Mouse EST Project 1999  
Unpublished (1999)  
Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL : contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:307801  
This read is a RESEQUENCE of a previously sequenced mouse clone  
This read has been verified (found to hit its original self in the  
correct orientation)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 339.

TITLE  
JOURNAL  
COMMENT

1. 500  
Location/Qualifiers

/organism="Mus musculus"  
/strain="Inbred CD-1"  
/db\_xref="taxon:10090"  
/clone="IMAGE:513953"  
/clone\_lib="Stratagene mouse testis (#937308)"  
/sex="males"  
/tissue\_type="testis"  
/dev\_stage="10-12 week old"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Organ: testis; Vector: pBluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector;  
-5' adaptor sequence: 5' CTCGAGTTTGTGTTTGTGTTT 3'";  
sequence: 145 a 114 c 121 g 120 t

FEATURES  
source

Query Match 35.3%; Score 333.6; DB 10; Length 500;  
Best Local Similarity 82.3%; Pred. No. 1.5e-66;  
Matches 408; Conservative 0; Mismatches 84; Indels 4; Gaps 2;  
QY 135 gaataagcactaagaagaatacgtatgttgaagcaattaaacatcccaatcttgtgaa 194  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2 GAATAAGCCTGCGGGAATCGTATGCTGAAG---TTGAAACACCCAACTCGTGAA 58  
QY 195 cctcatgaggtgttcaaggagaaagaaatgcattttagtttttgaatactgtatca 254  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 59 CCTCATGAGGTGTTCAAGAAGAAAGAGATGTCATCTAGTTTGTGACTGTGATCA 118  
QY 255 tacacttttaataagctgggaaagaaacccaatggaggtgtgctgagtgatcaaaag 314  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 119 CACACTGTTAAACAGCTGGAGAGAAACCCAAACGGAGTTTCTGATGAGTGATTAAAG 178  
QY 315 cgtattatgcaaacactcaagctcttaatttctgtcataacataactgtattcaag 374  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 179 TGTGCTATGGCAAAACCTTCAAGCCCTTACTTCTGTCACAGCACAAATTGATTCATCG 238  
QY 375 agatataaacctgaaatattcttaataactaagaagaagaataatcaagatttgtgaatt 434  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 239 GGATGTAACCTGAAACATCTCTAATAACCAAGCAAGGAGATGATAAGATTGTGACTT 298  
QY 435 cgggtttgcaaaattctgtattccagagagatgcttacaccgattatgtactacagatg 494  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 299 TGGATTTCAGCAATTCCTAATTCAGGAGACGCTACACAGACTATGTTGCCACAGGTG 358  
QY 495 gtaccgagctctgaaactctgtgtggagatactcaagtgtgttctcaatcatatg 554  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 359 GTACCGAGCCCTCGAACTTCTCGTGGGAGACGAAGTACGGTTTCCCTCTGTGACCTGTG 418  
QY 555 ggctattggtgtgttttttgcagagctctctgacagggccagccactgtggcctggaatac 614  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

BASE COUNT  
ORIGIN

145 a 114 c 121 g 120 t

Db 419 GGCCTCGCGTGTGTTTATGCAGAGACTTCTGACGTGTCAGCCACTCTGCGCGG-AAATA 477  
QY 615 agatgtggaccaactt 630  
||| ||||| ||||| |||||  
Db 478 CGAGCTGGACCACCTT 493  
RESULT 3  
AA061797 938 bp mRNA EST 03-FEB-1997  
LOCUS m134109.r1 Stratagene mouse testis (#937308) Mus musculus cDNA  
DEFINITION clone IMAGE:513953 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN  
KINASE KXIALRE (HUMAN);, mRNA sequence.  
AA061797  
ACCESSION AA061797.1 GI:1555606  
VERSION  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 938)  
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterson, R.  
TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Marra M/Mouse EST Project  
Washington University School of MedicineP  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL : contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:307801  
Seq primer: -28ml3 rev1 ET from Amersham  
High quality sequence stop: 442.  
Location/Qualifiers  
1. 938  
/organism="Mus musculus"  
/strain="Inbred CD-1"  
/db\_xref="taxon:10090"  
/clone="IMAGE:513953"  
/clone\_lib="Stratagene mouse testis (#937308)"  
/sex="males"  
/tissue\_type="testis"  
/dev\_stage="10-12 week old"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Organ: testis; Vector: pBluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector;  
-5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor  
sequence: 5' CTCGAGTTTGTGTTTGTGTTT 3'".  
BASE COUNT 254 a 215 c 236 g 233 t  
ORIGIN  
Query Match 33.5%; Score 316.8; DB 10; Length 938;  
Best Local Similarity 82.9%; Pred. No. 1.1e-62;  
Matches 398; Conservative 0; Mismatches 77; Indels 5; Gaps 3;  
QY 135 gaataagcactaagaagaatacgtatgttgaagcaattaaacatcccaatcttgtgaa 194  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1 GAAATAGCCTCGGGAATCCGTATGCTGAAG---TTGAACACCCAACTCGTGAA 57  
QY 195 cctcatgaggtgttccagagaaagaaatgcattttagtttttgaatactgtatca 254  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 58 CCTCATGAGGTGTTCAAGAAGAAAGAGATGTCATCTAGTTTGTGACTGTGATCA 117  
QY 255 tacacttttaataagctggagaaagaaacccaatggaggtgtgctgagtgatgatacaag 314

Db	118	CACACTGTTAAACGAGCTGGAGAGAAACCCAAACGAGTCTTCTGATGGAGTGAATTAAGAAG	177
Qy	315	cgtattatggcaaacacttcaagctcttaattctctgtcatatacataaactgtatttcacag	374
Db	178	TGTGCTATGGCAACCCCTTCAAGGCCCTTAACCTCTGTCACAAACACAATTTGATTATCTCG	237
Qy	375	agataaaaaacctgaaaaattcttaataaactaagcaaggaataatacaagattttagactt	434
Db	238	GGATGTAAAACCTTGAACACATCCTTAATAACCAAGCAAGGATGATAAGCAATTTGTGACTT	297
Qy	435	cggggtgtgcacaattcttgattccagagaatcctcacagattatgtagctacagatg	494
Db	298	TGGATTGTGCAGGAATTCCTAAATTCACGAGACGCCCTACACAGACTATGTGGCCACCAAGTG	357
Qy	495	gtaccgagctctcgaactctgtgtggagataactcagtatggttctcagtcgatatg	554
Db	358	GTACCGAGACCCCGAANCYTCCTGTGGAGACACGAAGTACGGTTCTCTGTGTAGACGTGTG	417
Qy	555	agctattgtgtgttttgcagagctcctgcagggccagccactgtggcctggaaaaatc	614
Db	418	GGCCGTCGGCTGTGTTTTTGGACG - CTCCTCAAGGGTCAAG - CACTCTGGCCGGGAAACCC	475

RESULT	4
LOCUS	AW106692
DEFINITION	um32d03.y1 sugano mouse kidney mkoa Mus musculus cDNA clone IMAGE:2236229 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE K2ALRE (HUMAN); mRNA sequence.
LOCUS	AW106692 731 bp mRNA EST 20-OCT-1999
DEFINITION	um32d03.y1 sugano mouse kidney mkoa Mus musculus cDNA clone IMAGE:2236229 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE K2ALRE (HUMAN); mRNA sequence.

be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Cloning primers for sequencing: 5' end primer CTTCTGCTCTAAAGTGG and 3' end primer CGACCTGAGCTGACACA.

BASE COUNT	ORIGIN
202 a 179 c	191 g 157 t
2 others	

PRIMER	CGACCTGCAGCTCGAGCACA."	202 a	179 c	191 g	157 t	2 others
BASE COUNT						
ORIGIN						

Query Match	27.4%	Score 258.6	DB 10	Length 731	
Best Local Similarity	66.9%	Pred. No. 2.5e-49			
Matches 382	Conservative	0	Mismatches 186	Indels 3	Gaps 1
QY	5	aaaagtatgaaaaattagctaaagctgagaagggtctctatggggtgtattcaaatgca	64		
Db	156	AAATATCAAAAAAATTTGGAACATTTGGAGAAGGCTCCTATGGGTAGTGTTCAGTGC	215		
QY	65	gaaacaaacctctggacaagtagctgtgtataaaatttgggaactctggaagtatc	124		
Db	216	GAACAGGACACCGGTCAGATCGTGGCCATCAAGAGGTTTCTGGAACCCGAAGATGACC	275		
QY	125	ctgtgttaagaataatgacctaaagaaatcagtatgttgaagaacaataaaacatccaa	184		
Db	276	CTGTCTATAAGAAAATCGCCCTTCAGAAATTCGCATGCTCAGCACTCAGCACC	335		
QY	185	atcttgtgaacctcatcgagggtgttcaggagaaaaaggaaatgcatttagttttgaat	244		
Db	336	ACCTGGTCAACCTCTGGAAAGTCTTCGGAGGAAGCGGAGCGTTCACTGGTGTTCGAGT	395		
QY	245	actgtgatcatcacactttaaattgagctggaagaaacccaattggagttgctgatgag	304		
Db	396	ACTGGCAGCACACGGTGCTTCACAGAGCTGGATCGGTATCAGAGGGGGTACCAGAGCCTC	455		
QY	305	tgtatcaaaagcgttatatggcaaacacttcaagctcttaattctgtctatatacataact	364		
Db	456	TCGTGAACAACATAACTTGGCAGACACTGCAGGCTGTAAATTTCTGCCATAAACATAACT	515		
QY	365	gtattcacaagagataaaaaacctgaaatattctaatcaactaaagcaaggaaataatcaaga	424		
Db	516	GCATACACAGACAGCTGAAGCGCGAAATAATTCTCATCCCAACACAGTCAGCCATTAAAGC	575		
QY	425	tttgtgacttcggggtttgcacaaattcttgatt---ccaggagatgctcaaccattatg	481		
Db	576	TCTGTGACTTTGGGGTCCGACGGCTCCTACTGGACCANGTGACTACTACACAGACTACN	635		
QY	482	tagcttaagagatggttacccaagctctgaactcttctgtggagatactcaagtgttctt	541		
Db	636	GTGCCCAACCGCTGGTACTGCTCACCCGAGCTGCTAGTTTGGAGACAGCGAGTATTGGTCCC	695		
QY	542	caagtcatatagggctattggttgtttt	572		
Db	696	CTGTAGATGCTCGGCAATTCGGCTGTGTGTT	726		

FEATURES	high quality sequence stop: 433. Location/Qualifiers 1..731	RESULT 5 AA286088 LOCUS DEFINITION AA286088 718 bp mRNA vc33b05.r1 Barstead MPLRBI Mus musculus cDNA clone IMAGE:776337 5' similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE KIALRE (HUMAN );, mRNA sequence. ACCESSION VERSION KEYWORDS SOURCE house mouse. Mus musculus ORGANISM	09-APR-1997 EST Mus musculus cDNA clone IMAGE:776337 5' SERINE/THREONINE-PROTEIN KINASE KIALRE (HUMAN );, mRNA sequence. AA286088 AA286088.1 GI:1932198 EST. house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 718) REFERENCE AUTHORS Marras,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
----------	---	--	---





); mRNA sequence.		ACCESSION		AI508835		VERSION		AI508835.1		GI:4407740	
house mouse.		KEYWORDS		EST.		SOURCE		Mus musculus		house mouse.	
ORGANISM		REFERENCE		AUTHORS		TITLE		JOURNAL		COMMENT	
3;		20.4%;		Score 192.4;		DB 12;		Length 1535;		Query Match	
Best Local Similarity		54.6%;		Pred. No. 3.9e-34;		Mismatches 366;		Indels 12;		Gaps	
Matches 454;		Conservative 0;		Mismatches 366;		Indels 12;		Gaps		1 atggaagaagtataaaaaattagctaaagctgagaaggtctttagtggtgtattcaaa 60	
Qy		1 atggaagaagtataaaaaattagctaaagctgagaaggtctttagtggtgtattcaaa 60		Db		306 ATGAATAAATTTGAGATCTTGGGGTGTAGGTGAAGAGAGAGCTATGAGTGTACTTAAA 365		Qy		61 tgcagaacaacacctctggacaagtagctgtgttaaaaaatttggaaatcgaagat 120	
Db		366 TGCAGACACAGGAACACATGAAATTTGGCGATCAAGAAATTCAGAGACAGTGAAGAA 425		Qy		121 gatctgtgttaagaataacacataagagaataacgtatgttgaagcaattaaacat 180		Db		426 AATGAAGAAGTCAAAAGAACGACTTACGAGAGCTTAAATGCTTCGGAAGCTCAAGCAG 485	
Qy		181 ccaaatcttgtgaacctcatcgaggtgttcaggagagaaaaaattgcttagtttt 240		Db		486 GAAACATTTGGAGTTGAAGGAAGCATTTGTCGGAGGGGAAAGTGTACTTGGTGT 545		Qy		241 gaatactgtcatcacacttttaaatgagctgagaagaaacccaaatggagttgctgat 300	
Db		546 GAGTATGTTGAAAAAATATGCTCGAATTTCTGGAGAAATCCCAATGGAATTCACCT 605		Qy		301 ggaagtatacaaaagcgtattatggaacacacttcaagctcttaatttctgcatacat 360		Db		606 GAGAAAGTAAAAAGCTACATCTATCAGCTAATCAAGCTTAACTGTGCTGCCATAGAAC 665	
Qy		361 aactgtattcacagagatataaaacctgaaatattcttaataactaaagaaggaataatc 420		Db		725 GATATTGTCCTCGAGATATAAACCCAGAAATCTCTTAATCAGCCAGACAATGATGCTTA 725		Qy		421 agatttgtgacttgggtttgcacaaattctgattccagg-----agatgctcacacc 474	
Db		726 AAAGTGTGACTTTGGTTTGTCTGCTAATCTGTCAAGAGCAATANTGCTAATTACACA 785		Qy		475 gattatgtagctagagatggtaccagctcctgaacttcttggggagatactcaat 534		Db		786 GAGTACGTTGCCACAGATGTATCGGTCCCGAGAACTCTTACTTGGCG---CTCCCTAT 842	
Qy		535 ggtcttcagtcgatatatggctattgtgttttttcgagagctctcctgacagggccag 594		Db		902 GGAAGTCCGTGGACATGCTGTCGTTGGCTGTATTTCTTGGGAGGCTTAGCGATGGACAG 902		Qy		595 ccactgtggcctggaaaaatcaagatgtggacaaactttatctgataatcagaacactagga 654	
Db		902 CTTTATTTCTCGGAAGTGAATTTGACCACTTTTACTATTTCAGAAAGTGTGAGGA 962		Qy		655 aaatlaatcccaagacatcaatcattttaaagaacaggggttttccatggcatcagt 714		Db		963 CCATCTCCATCTCAGCAGATGAAGCTTTTCTACATACTCTGCTTCCATGGGTCGG 1022	
Qy		715 ataactgagccagagacatggaactcttggagaa---agttctcagatgttcatcct 771		Db		1023 TTTCAGCTCTTAAACCATCTCAGCTCTCTGGAAAGAAGATACCTTGAATTTGAAATG 1082		Qy		772 gtggctctgaacttcatgaaggggtgtcgaagatgaatccagatgacagat 823	
Db		1083 GTTCTACTTGACCTAATGAAGAAATTTACTGAAGTTGGACCCAGCTGACAGAT 1134		RESULT		9		AI508835		727 bp	
LOCUS		mRNA		EST		12-MAR-1999		AI508835		Mus musculus cDNA clone IMAGE:776337 5'	
DEFINITION		similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE KIALRE (HUMAN									

house mouse.		BASE COUNT		200 a 169 c 191 g 158 t		ORIGIN	
Query Match		19.7%;		Score 186.6;		DB 10;	
Best Local Similarity		64.5%;		Pred. No. 8.4e-33;		Length 727;	
Matches 287;		Conservative 0;		Mismatches 157;		Indels 1;	
Gaps		1;					
Qy		1 atggaagaagtataaaaaattagctaaagctgagaaggtctttagtggtgtattcaaa 60		Db		110 ATGAAAAATATGAAAAAATTTGGAAGAGATGGGAAGCGTCTATGGGTAGTGTCAAG 169	
Qy		61 tgcagaacaacacctctggacaagtagctgttataaaaaatttggaaatcgaagat 120		Db		170 TGCAGAAACAGGACACGGGTTCAGATCGTGGCCATCAAGAGGTTCCTGGAACCCGAAGAT 229	
Qy		121 gatcctgttataagaataacacataagagaataacgtatgttgaagcaattaaacat 180		Db		230 GACCTCTCATAAAGAAATCGCCCTTCGAGAAATCCGCATGCTCAAGCAACTCAAGCAC 289	
Qy		181 ccaaatcttgtgaacctcatcgaggtgttccagagaaaaaagaaatgaatttttt 240		Db		290 CCCAACCTGGTCAACCTCTCGAAAGTCTCCGAGGAGCGGAGGCTTACCTTGGTGTTC 349	

); mRNA sequence.  
AI508835  
AI508835.1 GI:4407740  
EST.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 727)  
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
Underwood, K., Steptoe, M., Rheising, B., Allen, M., Bowers, Y., Person,  
B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter,  
E., Kohn, S., Shin, T., Jackson, J., Cardenas, M., McCann, R.,  
Waterston, R. and Wilson, R.  
The WashU-NCI Mouse EST Project 1999  
Unpublished (1999)  
Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:469193  
This read is a RESEQUENCE of a previously sequenced mouse clone  
correct orientation)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 399.  
Location/Qualifiers  
1. .727  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/clone="IMAGE:776337"  
/clone\_lib="Barstead MPLRB1"  
/sex="mixed"  
/tissue\_type="Kidney"  
/dev\_stage="6 weeks"  
/lab\_host="DH10B"  
/note="Vector: p773D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: EcoRI; Site\_2: NotI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTAGCAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
[CATGATTCGCTACC], digested with Not I and cloned into the  
Not I and Eco RI sites of the modified p77T3 vector.  
Library constructed by Bob Barstead."

```

QY 241 gaatactgtgatcatacaacttttaataatgagctggaagaaccccaaatggagttgctgat 300
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 GAGTACTGGCAGCACACCGTCTTCACAGCNTGGATCGGTATCAAGGGGGGTACACAG 409

QY 301 ggaagtatcaaaagcgtattatgcaaacacttcaagctcttaatttctgtcatatacat 360
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 CTNCTCGTGAAGA-CATNNACTTGCAGACACTGCAGGCTGTNAATTTCTGGCATTAAACA 468

QY 361 aactgtattcacagagatataaaacctgaaatatttctaacttaactaagcaagggaataatc 420
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 TACTGCATACCAGGACGTCGAAAGCCGGAATAATTTCTATCAACAAACAGTCNAGCATT 528

QY 421 aagatttggacttgggtttgcac 445
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 AAGCCCTGTATNNTTTGGGTTTCGAC 553

RESULT 10
LOCUS AV672492 582 bp mRNA EST 05-OCT-2000
DEFINITION AV672492 Nori Satoh unpublished cDNA library Ciona intestinalis
ACCESSION AV672492
VERSION AV672492.1 GI:10110491
KEYWORDS EST.
SOURCE Ciona intestinalis.
ORGANISM Ciona intestinalis.
REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
AUTHORS Phlebobranchia; Clonidae; Ciona.
TITLE 1 (bases 1 to 582)
JOURNAL Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
COMMENT Expressed genes in Ciona intestinalis
Unpublished (2000)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satohe@scidtan.zool.kyoto-u.ac.jp.

FEATURES
    source
        Location/Qualifiers
            1..582
                /organism="Ciona intestinalis"
                /db_xref="taxon:7719"
                /clone="citb5a12"
                /clone_lib="Nori Satoh unpublished cDNA library"
                /tissue_type="whole animal"
                /dev_stage="tailbud"
BASE COUNT 178 a 100 c 143 g 159 t 2 others
ORIGIN
    Query Match 18.5%; Score 174.8; DB 10; Length 582;
    Best Local Similarity 60.3%; Pred. No. 4.3e-30;
    Matches 323; Conservative 0; Mismatches 209; Indels 4; Gaps 2;

QY 1 atggaagaatgaaaaattagctaaagactggagaggggtcttattgggttgattcaaa 60
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 47 ATGGAGAAATATGACACCTTGGCATGGGCTGGGTGAAGGCTCTATGGGATGTTATGAAA 106

QY 61 tgcagaacaaacacctgtggacaagttagctgttataaaatttggaaatctgaagat 120
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 TGTAAACATAAGGAAAGTGGGCAAAATGTTGCCATTAAAGAGTTTCTGGAAGTGAAGAT 166

QY 121 galcctgttgttaaaaaatagc-actaagagaaatcgtatgttgaaagcaattaaaaa 179
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 GACAACATGTTGAAAGATTGGCATGCNGTGAAGTTTCGAATGTTTGAGAAAACATCCATCA 226

QY 180 tccaaatcttgaacctcatcaggtgttcaggagaaaaagaaaaatgcattagtttt 239
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 CGAAAACCTCGTCAACTTAATTAAGTTTTCGTCGACGTAAGCGTCTTTATCTGCTCTT 286

QY 240 tgaatactgtgatcatacaacttttaataatgagctggaagaaccccaaatggagttgctga 299

```

```

Db 287 TGAATTTGTGCACACACCGTGTGGATGACCTTGANAAATATCTTAACGGATTAAACGA 346
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 300 tggagtgatcaaaagcgtattatgcaaacacttcaagctcttcaatttctgtcatataca 359
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 347 GATGACAGTTTGAAAAATTCCTGTGGCAAGTTCTTAGAGGAGTAGAATTTCTGTACAGCCA 406
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 360 taactgtattcacagagatataaaacctgaaatatttctaactaactaagcaagggaataat 419
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 TAATATCATACACAGGATATCAAACTTGAATAATATTTAAACTCAGCATCGGGGTCAT 466
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 420 caagatttggacttgggtttgcac---aaattctgattccagagatgctcacaccga 476
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 467 CAAACTATGTGACTTTGGTTTTCGTCGGACATTAGCTGCTCTGTGAGGTTTATACGA 526
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 477 ttatgtactacagagatggtaccgagctcctgaactcttctgtggagatactcagt 532
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 527 TTATGTGGCCACTAGGTGTATCTGTCGCCGAGCTGTTGGTTCGGAGACACCAAT 582
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 11
LOCUS AQ953706 665 bp DNA GSS 27-JAN-2000
DEFINITION Sheared DNA-53L19.TF Sheared DNA Trypanosoma brucei genomic clone
ACCESSION AQ953706
VERSION AQ953706.1 GI:6776971
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 665)
AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
Gerard,C., Leech,V., de Jong,P., Ollu,E., Melville,S., Donelson,J.,
Fraser,C. and Adams,M.
Determination of clone end sequences from Trypanosoma brucei GUTat
10.1 sheared DNA library
Unpublished (1999)
Other_GSSs: Sheared DNA-53L19.TF
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: nelsayed@tigr.org
Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
DNA library constructed at TIGR. Clones will be available for
distribution through ATCC. Sheared DNA end sequences search page:
http://www.tigr.org/tldb/mdb/tbdb/.
Seq primer: M13-Forward
Class: shotgun.

```

```

FEATURES
    source
        Location/Qualifiers
            1..665
                /organism="Trypanosoma brucei"
                /strain="TREU927/4 GUTat 10.1"
                /db_xref="taxon:5691"
                /clone="Sheared DNA-53L19"
                /clone_lib="Sheared DNA"
                /note="Vector: pUC18; Site 1; Small; Constructed at The
                Institute for Genomic Research (TIGR), Rockville, MD.
                Genomic DNA isolated from a cloned population of
                Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
                sheared to give a tight size distribution (approx 2 kb).
                The v + i method used for the library construction is
                described in detail in Smith, H.O. and Venter, J.C.
                (Making small insert libraries for whole genome shotgun
                sequencing projects. In Genome Sequencing: A Practical
                Approach, eds. M. Vaudin and B. Borell, Oxford University
                Press, 1999)."
BASE COUNT 183 a 122 c 190 g 170 t
ORIGIN

```



Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=4t2-QV3-NN0099-020  
800-286-e114t3=2000-08-02&t4=1)

Seq primer: puc 18 forward  
High quality sequence start: 20  
High quality sequence stop: 399.  
Location/Qualifiers

FEATURES  
source

1..399  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN0099"  
/dev\_stage="Adult"  
/note="Organ: nervous\_normal; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
93 a 97 c 111 g 98 t

BASE COUNT  
ORIGIN

Query Match 18.1%; Score 170.6; DB 10; Length 399;  
Best Local Similarity 65.2%; Pred. No. 4e-29; Mismatches 134; Indels 0; Gaps 0;  
Matches 251; Conservative 0;

QY 461 gagatgcctacacgattatgtacgtacgagatgtaccagctcctcctgaactcttctgtg 520  
Db 12 GTGACTACTATGACAGACTAGTGGCTACCGGTGGTACCGCTCCCTCAGCTGCTGGTGG 71  
QY 521 gagatcactgaatgttctcagtcgatatatgagctattgtgtgttttttcagagc 580  
Db 72 GGGACACGACGTACGAGCCCGCCCGGTGGATGTTGGGCAATTTGGCTGTCTCTCTGAGC 131  
QY 581 tcttgacagccagccactgtgctggaataatcagatgtgacacactttatctgataa 640  
Db 132 TGCTGTGACGAGTGCCTCTGTGGCCAGGAAATCGGATGTGGATCATCTGATTA 191  
QY 641 tcagaacactaggaataatcccaagacatcaatcctttaaagtaacgggtttt 700  
Db 192 GGAAGACCTTGGGGATCTCATTTCTAGGCACGCAAGTGTTCACGAAATCAGTACT 251  
QY 701 tccatggcatcagtatacctgagccagagacatggaactcttgaggagaaagtctcag 760  
Db 252 TCAGTGGAGTGAATAATTCAGACCCCTGAAGATATGGAACCACTTGAATTAATTTCCCA 311  
QY 761 atgttcactcgtggtcctgaacttcacgaagggtgtctcgaagatgaatccagatgaca 820  
Db 312 ACATCTCTTATCCTGCCCTGGGGCTCTAAAGGGCTGTCTCCACATGGACCTACTGAAA 371  
QY 821 gattaacctgttcccaactcctgga 845  
Db 372 GGCTGACATGTGAACAGCTGTTGCA 396

RESULT 13  
LOCUS

BE700389 401 bp mRNA EST 12-SEP-2000  
DEFINITION QV3-NN0099-100800-267-f09 NN0099 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE700389  
VERSION BE700389.1 GI:10088131  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 401)

Query Match 18.5%; Score 174.4; DB 13; Length 665;  
Best Local Similarity 56.4%; Pred. No. 5.3e-30;  
Matches 325; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

QY 1 atggaagatgataaaattagctaaagactggagaaggtcttattgggtttgtattcaaa 60  
Db 90 ATGATCGGTACGAACACTTGGTATGCTTGGGAGGGAGCTTATGCTGTAGTGTGAAG 149  
QY 61 tgcagaaacaaacctctggacaagtagtagctgtttaaataatttgggaactcgaagat 120  
Db 150 GCGCGCACCGCGGCATATCAGCATGCTGTATTAAAGATACAAACAGGCTGAAGAT 209  
QY 121 gatcctgttgttaagaaaatagcactaagaagaaatcgtattgttgaagcaattaaacat 180  
Db 210 GACGATCAGCTCCGCAAAACGTCGCTTCGGGAGGTACGTGTGCTGAAGCAGTTACGCCAC 269  
QY 181 ccaaatcttgtgaacctcaggtgttcaggagaaaaagaaaatgcatttagtttt 240  
Db 270 CTAACGTACAGCCCTGTTAGATGATTTCGCCGCGATGGGAAGCTTTATCTGTTTTT 329  
QY 241 gaactactgtatcacactctttaaataagctggaagaaagaaacccaatggagtgtgat 300  
Db 330 GAGTATGTTGAGAACACGATTTCTGCAGCTGATTGAGGAAAGAAAGTATGGACATTCACCC 389  
QY 301 ggagtgatcaaaagcgtattatggcaaaccttcaagctcttaatttctgtcatatcat 360  
Db 390 GATGAGGTACCGCGGTACACCTTTCAGTTATTGACGGTGTGAGTTACTGCCATGCCACAC 449  
QY 361 aactgtattcacagagatataaaacctgaaataattctataactaagaagaataatc 420  
Db 450 AATATCAITTCACCGCGATGTAAAGCCAGAAACATCCTAGTATCAGGGAGCGGCTCTG 509  
QY 421 aagatttgacttcggttttgcaaaattctgattccagagagatgcctacacagcattat 480  
Db 510 AAGTTGTGCGACTTGTGTTTCGCCAGGAGTGTGCTGTCGGGGAAATATACGGAATAT 569  
QY 481 gtatgtacagatggtaccagctcctgaactcttgttgagagatactcagatggttct 540  
Db 570 GTTGCACACTCGTGTGATCGTCGACCGGAGTTACTAGTTGGTGATGATCATATATGCGCAA 629  
QY 541 tcaatgcataatgggtctattggtgttttttgc 576  
Db 630 CGGTTGATGTTGGCTATGTTGTTGCGTTTTTCA 665

RESULT 12  
BE700387

LOCUS BE700387 399 bp mRNA EST 12-SEP-2000  
DEFINITION QV3-NN0099-020800-286-e11 NN0099 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE700387  
VERSION BE700387.1 GI:10088129  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 399)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL 20202663  
MEDLINE  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

## AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Colman, G. H., Carvalho, A. F., Matsumura, A., Bala, G. S., Simpson, D. H., Brunschein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, S. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. *Proc. Natl. Acad. Sci. U.S.A.* 97 (7), 3491-3496 (2000).

02682003  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genom  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&td=QV3-NN0099-1>)  
800-267-f09&t3=2000-08-10&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 400.

FEATURES  
source

```

location/qualifiers
1. .401
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NN0095"
/dev_stage="Adult"
/note="Organ: nervous_normal; Vector: puc18; Site.1: Sma1;
Site.2: Sma2; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
94 a 98 c 108 g 101 t

```

Query Match	17.5%	Score 165.6;	DB 10;	Length 401;
Best Local Similarity	65.0%;	Pred. No. 5,6e-28;		
Matches 260;	Conservative 0;	Mismatches 139;	Indels 1;	Gaps
QY	461	gagatgcctacacgagatgatgaactacagatggttaocgagctctgaactctctgtg	520	
Db	2	GTGACTACTATACAGACTACCTGGCTTACCAGTGTGATCCGCTCCCTGAGTGTGTGGTG	61	
QY	521	gagatactcagtatggtcttcca-gtcgatatatggctatggttggttttttgcagag	579	
Db	62	GGGACAGCAGTAGCGCCCCCGGGTGGATTTGGGCAATGGCTGTGCTTTGTGTGAG	121	
QY	580	ctcttcacagcaggcagcactgtggcctggaaatcagatggtgaccaactttactgata	639	
Db	122	CTGCTGTGACAGAGTGCCTCTGTGGCCAGGAAATCGGATGGATCGACTGTATCTGATT	181	
QY	640	atcagaacactaggaaaattaaatcccaagacataaatcaatcttttaaagttaacggggtt	699	
Db	182	AGGAAGACCTTGGGGGATCTCATTTCTTAGGCACCCAGCAGTGTTTAGCAGAAATCAGTAC	241	
QY	700	ttccatggcatcagtatacctgagccaggaagacatggaaactcttgaggaaaagtctcca	759	
Db	242	TTCAGTGGAGTGAANAATTCAGACCTTGAAGATATGGAACCACTTTGAATTAAAATTC	301	
QY	760	gatgttcactcctgtgctctgaacttcgatgaagggtgtctgaagatgaatccagatgac	819	
Db	302	AACATCTCTATCTGCTGGGCTCTTAANGGGCTGTCTCCACATGGACCCCTACTGAA	361	
QY	820	agataaactgttcccaactcctgagagctcctacttgg	859	
Db	362	AGCTGACATGTGAACAGCTGTGTGATCACCACATATTTTG	401	

RESULT	14
AI385990	
LOCUS	
DEFINITION	

AI385990 495 bp mRNA EST 27-JAN-1999  
ml50h06.y1 Stratagene mouse testis (#937308) Mus musculus cDNA  
clone IMAGE:515483 5' similar to TR:Q92772 Q92772 P56 KIAMRE  
PROTEIN KINASE. ;, mRNA sequence.

AI385990  
 ACCESSION  
 VERSION  
 AI385990.1  
 KEYWORDS  
 EST.  
 SOURCE  
 house mouse.  
 ORGANISM  
 Mus musculus

REFERENCE  
AUTHORS

1 mammalia; Eutheria: Rodentia: Sciurognathi: Muridae; Murinae; Mus.  
 1 (bases 1 to 495)  
 Marta, M., Hillier L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
 Underwood, K., Strepto, M., Theising, B., Allen, N., Bowers, Y., Person  
 E., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter  
 E., Kohn, S., Shih, T., Jackson, V., Cardenas, M., McCann, R.,  
 Waterston, R. and Wilson, R.

TITLE	JOURNAL	COMMENT

Contact: Marra M/WashU-NCI Mouse EST project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 MGI:309331

This read is a RESEQUENCE of a previously sequenced mouse clone  
 This read has been verified (found to hit its original self in the  
 correct orientation)  
 Possible reversed clone: similarity on wrong strand  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 358.

FEATURES	
SOURCE	

```

location/qualifiers
1. .495
/organism="Mus musculus"
/strain="Inbred CD-1"
/db_xref="taxon:10090"
/clone_image:515483
/clone_lib="Stratagene mouse testis (#837308)"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"
/notes="Organ: testis; Vector: pBluescript SK-
ECORI. Site-2: XhoI; Cloned unidirectionally.
Oligo dr. Average insert size: 1.0 kb; Uni-ZAP
-5' adaptor sequence: 5' GAATTCGCACAG 3' -3'
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
151 a 100 c 110 g 129 t 5 others

```

Query Match	16.8%	Score 158.4	DB 10	Length 495
Best Local Similarity	61.0%	Pred. No. 2.5e-26		
Matches 271	Conservative 0	Mismatches 170	Indels 3	Gaps 1
129	tgttaagaataatgacctaagaagaaatcgtattgtgaagcaattaaaaacatccaatct	188		
18	TGTCACAAATTTGCACGAGAGAATAAAGTTTCTAAAGCAATTTGGCATGAAAACCT	77		
189	tgtgaacctcatcgagtggttcaggagaaaaagaaatgcatttagtttttgaatactg	248		
78	GGTCAATTTGATTGAAGCTTTTATAGACAAAACAGAAATTCATTTGGTATTTGGGTTTAT	137		
249	tgaatcatacattttaatgactggagaaagaaacccaatggagtgtgctgatggagtgat	308		
138	TGACCACACGGCTCTAGTAGGCTGCACACTACTCTCAGGACTTAGAGAGTAAGCGCT	197		

ORIGIN

Query Match 16.4%; Score 155.4; DB 10; Length 384;  
Best Local Similarity 62.7%; Pred. NO. 1.2e-25;  
Matches 240; Conservative 0; Mismatches 143; Indels 0; Caps 0;

QY 287 atggagttctgatgagatgacaaagcggtattatggaacacacttcaagctcttaatt 346  
Db 2 AAGGATGTCCTGATGTTTGGATTAAAAAACTATTTATCAATTATGTTAGCTATAAGAT 61  
QY 347 tctgtcatatatacaactgtatttcaacagagataataaacctgaaaaattcttaacta 406  
Db 62 ATTCTCATAAACATAAATTTGATGCATAGAGATGTTAAACACAGAAAATATTTTAAACAA 121  
QY 407 agcaaggaataaacaagatttctgaacttgggttgcacaaattcttgattccagagagatg 466  
Db 122 GTAACGATGTTGTCAAAATAGCTGATTTTGGTTTGAAGAGTTGTCAATACAAAATGATT 181  
QY 467 cctacacccgattatgtactacagagatggtacacgagctccctgaacttcttggggagata 526  
Db 182 TATATACAGATTATGTTGCAACTCGTGGGTATAGATGCCCTGAATATTATTAGTTGGTGATA 241  
QY 527 ctcaagtattgtcttcagtcgatataatgggctattggttgggttttgcagagctcttga 586  
Db 242 TTGAATATGGGGTTAGTGTGTGATATTNGGCAATTTGGATGTGTATAGCTGAAATGGTTA 301  
QY 587 caggccagccactgtggcctgaaaaatcagatgtgacccaactttatctgataatacagaa 646  
Db 302 GTGGTGATGCAATATGGCCTGGTAGAAGTGTGATNAACTTTTTTTTGTGATATTTAAA 361  
QY 647 cactagggaataataatcccaaga 669  
Db 362 CATTTGGGTTCACCTCTGCAAGA 384

Search completed: May 2, 2002, 17:06:45  
Job time: 11057 sec

QY 309 caaaagcgtattatggcaaacacttcaagctcttaattctgtcatatatacaactgtat 368  
Db 198 GAGAAAGTACCTCTTCCAGATCCCTTCGAGCCATTGAGTACTGTCATAATAATATATAT 257  
QY 369 tcacagagatataaaacactgaaaattcttaataactaaagcagagatttg 428  
Db 258 CCATCGAGATATAAAGCCTGAGAAATATTTAGTCTCCAGCTCAGGAATTTACGAAGCTGTG 317  
QY 429 tgaacttcgggtttgcacaaattct--gattccagagatgctctacacccgattatgagc 485  
Db 318 TGATTTTGGGTTTGGCAGGACACTAGCAGCTCTCTGGAGACGTTTACACAGACTAGCTGGC 377  
QY 486 tacgagatggtacccagctcctgaacttcttggggagatactcagtatggttcttcagt 545  
Db 378 CACAGCGTGTACAGAGCTCCAGAGCTGGTGTGAAAGACACCTNCTACGGANAGCCAGT 437  
QY 546 cgatatatggctattggttgt 569  
Db 438 GGACATCNTGCTTTNGGCTGTAT 461

RESULT 15  
BE223187 384 bp mRNA EST 09-MAY-2001  
LOCUS kp73q08.v1.TBN95TM-SSPH Strongyloides stercoralis cDNA 5' similar  
DEFINITION to WP.Y42A5A.4 CE20258 EUKARYOTIC PROTEIN KINASE DOMAIN :, mRNA  
sequence.

ACCESSION BE223187  
VERSION BE223187.1 GI:8927762  
KEYWORDS Strongyloides stercoralis.  
SOURCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
ORGANISM Panagrolaimoidea; Strongyloidea; Strongyloides.

REFERENCE 1 (bases 1 to 384)  
AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T.,  
Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y.,  
Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagarelis, V., R.,  
Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe,  
M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S.,  
Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and  
Wilson, R.

TITLE The Washington Univ. Nematode EST Project, 1999  
JOURNAL Unpublished (1999)  
COMMENT Contact: McCarter JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
The library was constructed by Dr. Thomas Nutman and colleagues of  
NIH (Nutman@nih.gov). DNA sequencing by: Washington  
University Genome Sequencing Center St. Louis.  
Seq primer: -40RP from Gibco  
High quality sequence stop: 340.  
Location/Qualifiers  
1..384  
/organism="Strongyloides stercoralis"  
/strain="Filariform larvae obtained from humans"  
/db\_xref="taxon:6248"  
/clone\_lib="TBN95TM-SSPH"  
/lab\_host="XL-1 Blue MRF" (Stratagene)  
/note="Vector: Lambda Uni-ZAP XR (Stratagene); Site\_1:  
EcoRI; Site\_2: XhoI; mRNA was purified from 4 x 10E5  
filariform larvae which had been isolated from infected  
humans. cDNA was constructed and, using adaptors, was  
cloned unidirectionally into the vector from the EcoRI  
site to the XhoI site. The library has an unamplified  
titer of 1.5 x 10E6 pfu/ml and an amplified, undiluted  
titer of 7 x 10E9 pfu/ml. The average insert size of the  
unamplified library is 975 bp (range, 500-1500)."

BASE COUNT 120 a 37 c 81 g 144 t 2 others

-----



```
Query Match      18.2%; Score 171.8; DB 2; Length 903;
Best Local Similarity 52.3%; Pred. No. 2.4e-39;
Matches 458; Conservative 0; Mismatches 402; Indels 15; Gaps 3;

QY 1 atggaagatgataaaattagctgaagactggagagggctcttatgggtgtattcaaa 60
Db 1 ATGGAGCAATATCAGAGGTTAGAGAAGATTGAGAGGAACCTTATGGAGTTGTTTATAAA 60
QY 61 tgcagaacaacacccctgcagacagtagtagctgttaaaaaatttggaaatctgaagat 120
Db 61 GCAAGGATCTTGAAGTGTACAAATTGTAGTCTTTAAGAAAAATCCGATTAGAAGCAGAA 120
QY 121 gatcctgtgttaagaataatagcactaagagaataatcgatgttgaagcaattaaacat 180
Db 121 GATGAGGGAGTCTCTAGTACAGCAATTCGTGAGATATCACTTTTGAAGACATGCCAAT 180
QY 181 ccaaatctgtgaacctcatcgaggtgttcaggagagaaaaagaaatgcattagttttt 240
Db 181 GATAATGTTGTAAGACTTTTGAATATATATTCATCAAGATCGCTTTATCTGTGTTTTT 240
QY 241 gaatactgtatcatacactctttaaataagctggaagaaagaaacccaatg-----gagtt 294
Db 241 GAATTTCTGTGATCTTGATTTAAAAAAGTATATGAATAGTATTCACAAAGGACATGATGCTT 300
QY 295 gctgatggatgatcaaaagcggtattatggcaaacacactcgaagctcttaattctgcat 354
Db 301 GGTGCAGAAATGATTAAAAAGTTTATGTCACACTTGTATCAGGTGTAAATATTTGTCAAT 360
QY 355 atacataactgtatcacagagataaaaacccctgaaatattcttaataactaagaagga 414
Db 361 TCTCATCTGATTCTCATCTGCTGACTTGAACACCAACAAAATCTTCTATAGATCGAGAAGGA 420
QY 415 ataatacaactgtgacttogggtttgca---caaatctgattccagagagatccctac 471
Db 421 ATATCTAAATATAGCAGATTTGGGCTTGCAAGGGCGTTTGGTGTCTCCATTCCTGGTTAT 480
QY 472 accgattatgtactcagagatggtaccagactcctgaactcttctggtgagatgactcag 531
Db 481 ACTCATGAAGTTGTTACACTTTGGTATCGTGTGCTCCAGAGGTTCTTTTAGGTGGTCGACAA 540
QY 532 tatggttcttcagtcgatatatgggtattggtgtttttgagagctcctgacagc 591
Db 541 TATGCACACGGCTGTATATATGAGCATTTGGATGATATTTTGCAGAAATGGCTACAAA 600
QY 592 cagccactgtyggcctggaaaatcagatgtggaccaactttatctgataatcagaacacta 651
Db 601 AAGCCATATTTCAGGTCATCTCGAAATTTGATGAATATTTAGAATATTTAGAATATTA 660
QY 652 ggaataatcccaagacatcaatcaatctttaaagtaaacgggtttttccatggcactc 711
Db 661 GGGACTCCAGATGAAATTTCTGGCTGGTATTTACATCTTATCCGGATTTTAAAGCAACT 720
QY 712 agtataacctgaccagaagacatggaacctcttgaggaaaagttctcagatgttcatcct 771
Db 721 TTTCACAAATGGTCACCA-----AAAAATCTTGAGAAATTAATTACAGAACTTGATAGT 774
QY 772 gtggtctgaacttcagaggggtgtctgaagatgaatccagatgacagatcaactcgt 831
Db 775 GATGGAATAGATTTATTACAGAAATGCTCTAGATATTTATCTGCTGCTGAACGTTATTAGCGCT 834
QY 832 tcccaactcctggagagctcactcttattgattttt 866
Db 835 AAAAAAGCTCTCGATCATCTTATTTTGTGATTTT 869
```

## RESULT 2

```
US-09-093-522-9
; Sequence 9, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leol, Edward B.
; APPLICANT: Thomas, Charles F.
```

```
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCOYSTIS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...900
; OTHER INFORMATION:
;
US-09-093-522-9
```

```
Query Match      18.2%; Score 171.8; DB 3; Length 903;
Best Local Similarity 52.3%; Pred. No. 2.4e-39;
Matches 458; Conservative 0; Mismatches 402; Indels 15; Gaps 3;
```

```
QY 1 atggaagatgataaaattagctgaagactggagagggctcttatgggtgtattcaaa 60
Db 1 ATGGAGCAATATCAGAGGTTAGAGAAGATTGAGAGGAACCTTATGGAGTTGTTTATAAA 60
QY 61 tgcagaacaacacccctgcagacagtagtagctgttaaaaaatttggaaatctgaagat 120
Db 61 GCAAGGATCTTGAAGTGTACAAATTGTAGTCTTTAAGAAAAATCCGATTAGAAGCAGAA 120
QY 121 gatcctgtgttaagaataatagcactaagagaataatcgatgttgaagcaattaaacat 180
Db 121 GATGAGGGAGTCTCTAGTACAGCAATTCGTGAGATATCACTTTTGAAGACATGCCAAT 180
QY 181 ccaaatctgtgaacctcatcgaggtgttcaggagagaaaaagaaatgcattagttttt 240
Db 181 GATAATGTTGTAAGACTTTTGAATATATATTCATCAAGATCGCTTTATCTGTGTTTTT 240
QY 241 gaatactgtatcatacactctttaaataagctggaagaaagaaacccaatg-----gagtt 294
Db 241 GAATTTCTGTGATCTTGATTTAAAAAAGTATATGAATAGTATTCACAAAGGACATGATGCTT 300
QY 295 gctgatggatgatcaaaagcggtattatggcaaacacactcgaagctcttaattctgcat 354
Db 301 GGTGCAGAAATGATTAAAAAGTTTATGTCACAACTTGTATCAGGTGTAAATATTTGTCAAT 360
```







```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
US-08-464-517-37

Query Match          14.2%; Score 134.6; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 8.5e-29;
Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

QY 1 atggaagatgataaaattagctaaagactgagaaaggtcttatgggttgatcaaa 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ATGCAGAAATACGAGAACTGGGAAGATTGGGAAGGACCTACGGAAGTGTCTCAAG 72

QY 61 tgcagaacaaacctctggacaagtagctgttaaaaaatttgggaatctgaagat 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCCRAAAACCGGAGACTCATGAGATCGTGGCTCTAAACCGGTGAGGCTGGATGACGAT 132

QY 121 gatctgtgttaagaaatagcactaagaaatactatgttgaagcaattaaacat 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GATGAGGCTGCGGAGTTCGCGCCCTCCGGGAGATCTGCTACTCAAGAGCTGAAGCAC 192

QY 181 ccaactctgtgaacctcatcgagggtgttcaggagaaagaaatgcatttatttt 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 ANGACATCTGACGGCTTCATGACGTCTGCACAGCGACAAGAGCTGACTTTGGTTTTT 252

QY 241 gaactgtgatcatcacacttttaaatgagctggaagaaaccccaatggagtgtgat 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GAATTCTGTACACAGGACCTGAAGAAGTATTTTCACAGTTGCAATGTGACCTCGATCCT 312

QY 301 ggagtgatcaaaagcgtattattggaacacttcaagctcttaatttctgatacat 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAGATTGTAAGTCAATCTCTTCCAGCTACTAAAGGGCTGGGATCTTGTCTATGCGCG 372

QY 361 aactgtatcacagagatataaaacctgaaataattcttaactaagcaaggaataatc 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 ATGTGCTACACAGGACCTGAAGCCCGCAGAACCTGCTAATAACAGGAATGGGAGCTG 432

QY 421 aagattgtgactcgggtttgcacaaa---ttctgattccaggagatgctcacaggat 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AAATTGGCTGATTTTGGCCCTGCTCGAGCCCTTTGGGATTCGCCGTGTTACTCAGCT 492

QY 478 tatgtagctacgagatggttacccagctcctgaacctctctgtggtggagatactcagtatgdt 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 GAGGTGGTCACACTGTGTACCGCCACCACCGGATGCTCTCTTTGGGGCCAAAGCTGACTCC 552

QY 538 tottcagtgatataatgggtctatgtgtgtgttttttcagagct---cctgacagccag 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 ACGTCCATCGACATGTGTGACCGGCTGCATCTTTTCGAGAGCTGGCCCAATGCTGGCGG 612

QY 595 ccactgtggctggaataatcagatgtggaccaacttttatgataatcagaacacattagga 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 CCTCTTTTCCCGCAATGATGTGCGATGACCAAGTTGAAGAGGATCTTCGAGTCTGCTGGG 672

QY 655 a 655
Db 673 A 673

RESULT 6
US-08-246-361A-37
; Sequence 37, Application US/08246361A
; Patent No. 598582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
```

```

; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
US-08-246-361A-37
```

```

Query Match          14.2%; Score 134.6; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 8.5e-29;
Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

QY 1 atggaagatgataaaattagctaaagactgagaaaggtcttatgggttgatcaaa 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ATGCAGAAATACGAGAACTGGGAAGATTGGGAAGGACCTACGGAAGTGTCTCAAG 72

QY 61 tgcagaacaaacctctggacaagtagctgttaaaaaatttgggaatctgaagat 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCCRAAAACCGGAGACTCATGAGATCGTGGCTCTAAACCGGTGAGGCTGGATGACGAT 132

QY 121 gatctgtgttaagaaatagcactaagaaatactatgttgaagcaattaaacat 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GATGAGGCTGCGGAGTTCGCGCCCTCCGGGAGATCTGCTACTCAAGAGCTGAAGCAC 192

QY 181 ccaaatctgtgaacctcatcgagggtgttcaggagaaagaaatgcatttatttt 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AAGAACATCGTTCAGGCTTCATGACGTCTGCACAGCGACAAGAGCTGACTTTGGTTTTT 252

QY 241 gaactgtgatcatcacacttttaaatgagctggaagaaaccccaatggagtgtgat 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GAATTCTGTACACAGGACCTGAAGAAGTATTTTCACAGTTGCAATGTGACCTCGATCCT 312

QY 301 ggagtgatcaaaagcgtattattggaacacttcaagctcttaatttctgatacat 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAGATTGTAAGTCAATCTCTTCCAGCTACTAAAGGGCTGGGATCTTGTCTATGCGCG 372

QY 361 aactgtatcacagagatataaaacctgaaataattcttaactaagcaaggaataatc 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 ANTGCTACACAGGACCTGAAGCCCGCAGAACCTGCTAATAACAGGAATGGGAGCTG 432
```



us-09-671-050-11.rni

Fri May 3 11:02:08 2002

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,997
; FILING DATE: 17-DEC-1992
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
; PCT-US93-09945-1

Query Match 14.2%; Score 134.6; DB 5; Length 1089;
Best Local Similarity 52.3%; Pred. No. 8.5e-29;
Matches 346; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

Qy 1 atggaagaatgataaaatagctaaagactgagaagaggtcttatggggtgtattcaaa 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ATGCGAAATACAGAAACTGGAAAGATTGGGGAAGCGACCTACGGAACCTGTTCAAG 72
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 61 tgcagaaacaaacctctggacaagtagctgttttaaaaaatttggaaatctgaagat 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCCAAAACCGGGAGACTCATGAGATCGTGGCTCTAAACGGGTGAGGCTGGATGACGAT 132
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 121 gatctgtgttaagaaataagcactaagagaaatagctatgttgaagcaatlaaaacat 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GATGAGGTTGGCCGAGTTCGCCCTCCGGAGATCTGCTACTCAAGAGCTGAAGCAC 192
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 181 ccaaatctgtgaacctcatcgaggtgttcaggagaaaaaggaatgattagtttt 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AAGAACATCGTCAGGCTTCATGACGCTCTGCACAGCGCAAGAAAGCTGACTTGGTTTTT 252
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 241 gaatactgtatcatacacttttaaatgagctggaagaaacccaaatggagttgctgat 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GAATTCGTGACAGGACCTGGAAGAGATATTTTCACAGTTGCAATGGTGACCTCGATCCT 312
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 301 ggagtgatcaaaagcgtattatgggcaaaccttcaagctcttaattctcatatcatat 360
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAGATTGTAAGTCATTCTCTTCCAGCTACTATAAAGGGCTGGGATTCGTGATAGCCGC 372
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 361 aactgtattcacagagatataaaacctgaaatattcttaataactaagcaaggaaatc 420
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 AATGCTGTACACAGGACCTGAAGCCCCAGAACCTGCTCTAATAAACAGGAATGGGAGCTG 432
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 421 aagattgtgacttcgggtttgcacaaa---ttctgattccaggagatgcctcacaccgat 477
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AAATGGCTGATTTTGGCCTGGCTCGAGGCTTTGGGATTCCTCCGCTGTTACTCAGCT 492
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 478 tatgtactacagatggtgaccgagctcctgaactcttcttggtggagatactcagtatggt 537
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 GAGTGTGTACACTGTGTACCGCCACCGGATGTCTCTTTGGGCGCAAGCTGTACTCC 552
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 538 tcttcagtcgatataggtattgattgttttgcagagct---ccctgacagggccag 594
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 ACGTCCATGCATGATGGTGTAGCCCGCTGCATCTTTGGCAGAGTGGCCCAATGCTGGGGG 612
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 595 ccaactgtggcgtgaaatcagatgtytgaaacacttttctgtataatcgaacacatgga 654
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 CTTCTTTTCCCGGCAATGATGCTGATGACCACTGTAAGAGGATCTTCCGAGCTGTGGGG 672
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 655 a 655
Db 673 A 673

RESULT 9
US-08-463-090B-3
; Sequence 3, Application US/08463090B
; Patent No. 5801015

; GENERAL INFORMATION:
; APPLICANT: Cottarel, Guillaume
; APPLICANT: Damagnez, Veronique
; APPLICANT: Draetta, Guilio
; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins from
; TITLE OF INVENTION: Human Pathogens, and Uses Related Thereto
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Elliot, LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,050B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV032.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-7000
; TELEFAX: (617) 832-1259
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 43..993
; US-08-463-090B-3

Query Match 13.9%; Score 131; DB 1; Length 1002;
Best Local Similarity 54.1%; Pred. No. 8.5e-28;
Matches 360; Conservative 0; Mismatches 290; Indels 15; Gaps 4;

Qy 8 agtatgaaattagctaaagactgagaaggtcttatggggtgtattcaaa---atgca 64
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 ATTATCAAGTCAGAAAAAGTCGGAGAAGGTACTTATGGGGTTGTTTATAAGCATTAG 118
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 65 gaaacaaacctctggacaagtagctgttataaaaaatttggaaatctgaagatgac 124
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 ATACCAAGCACAAATAATAGATTGTTGCATTAAAGAAAATTCGATTAGAAATCAGAAGATG 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 125 ctgttgttaagaaataagcactaagagaaatagctatgttgaagcaattaaacatcaa 184
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 AAGGTGTACCTAGTACCCCATTAGAGAATCTCGTTATTTAAAGAAATGAAGATGATA 238
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 185 atcttgtgaacctcatcgagggtgttcaggagaaaa---aaggaaaaatgcattagtttg 241
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 ATATCGTTCGATTATATGATATATTATTCATTCAGATTCTCATAAATATATTATTAGTATTG 298
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 242 aatactgtatcatcaccttttaaatgagctggaagaaaccccaaatggagttgctgatg 301
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 AATTTTGGATTTAGATTTAAGAAATATATGGAAGTATTCCTCAAGGAGTTGGACTAG 358
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 302 gag-----tgatcaaaagcgtatattgccaacacttcaagcttcaatttctcata 355
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 GGGCTAATATGATAAAAAAGATTATGAATCAATTAATTCGAGGATTATTAACATTGTCAAT 418
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 356 tacataactgtattcaagagatataaaacctgaaatatcttctaatcaactaagcaagaa 415
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 419 CTCATCGAGTTTACATCGTGAATTAACACCAAAATTTATTGATTGATAAAGAGGGA 478  
QY 416 taatcaagatttgacttogggttgacaaa---ttctgattccagagagatgcctaca 472  
Db 479 ATTTAAATTTAGCAGATTTGGATTAGCTCGAGCATTTGGAGTTCCATTAAAGACATATA 538  
QY 473 ccgattatgactacagagatggtaccagagctcctgaactcttctgtggagatactcagt 532  
Db 539 CTCATGAAGTTGTCATTATGATTCGAGCTCCGAAATCTTGTTAGGAGGGAACAAT 598  
QY 533 atggtcttcagtgatataatggtgctattggtgtttttgagagctcctgacagggcc 592  
Db 599 ATTCCACTGGGTAGATATGCTGTGTGATGATATTTGCTGAAATGCTAATAGGA 658  
QY 593 agccactggtcgggaaaaatcagatggtggacaaactttatctgataatcagacactag 652  
Db 659 AACCATTAATTCCTGGTGATTCAGAAATGTAGAAATTTTCGGAATTTCCGAATTTTAG 718  
QY 653 gaaaa 657  
Db 719 GAACA 723

RESULT 10  
US-08-463-090B-5  
; Sequence 5, Application US/08463090B  
; Patent No. 5801015  
; GENERAL INFORMATION:  
; APPLICANT: Cottarel, Guillaume  
; APPLICANT: Damagnez, Veronique  
; APPLICANT: Draetta, Guilio  
; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins from  
; TITLE OF INVENTION: Human Pathogens, and Uses Related Thereto  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley, Hoag & Eliot, LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,090B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MIV032.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1299  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1070 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 30..1058  
US-08-463-090B-5

Query Match 12.4%; Score 117.2; DB 1; Length 1070;  
Best Local Similarity 49.7%; Pred. No. 6.9e-24;  
Matches 327; Conservative 0; Mismatches 328; Indels 3; Gaps 1;

QY 7 aagatgaaaaattagctaagactggaagagggtcttattggggttgattcaaatcaga 66  
Db 96 AATTACACCAAGAAAAAAGTAGGGAAGGTACATATGCTGTGTGCTAGTTGGGTAAA 155  
QY 67 acaaaacctctggacaaagtagtagctctttaaaaaatttgggaactctgaagatgacct 126  
Db 156 CAAATCTCCCAACCAAGTCNAATTTGCCATCAAGAATAACAAACAGGATATTCAAGAT 215  
QY 127 gttgtaagaaaaatagacactaagaaatacgtatgtttgaagcaataaacaatcacaat 186  
Db 216 GGGTTGGATATGTCAGCATTTGAGAGAGTGAATAATTTGCAAGAAATGGAACATCCCAT 275  
QY 187 ctgtgaacctctcaggtgttcagggagaaaaagaaatgcattttagtttgaatac 246  
Db 276 GTTATTGAACCTAGTAGATGATTTTCACGACACAAATTAATTAATTTGGTATTGAATTT 335  
QY 247 tctgatacacactttttaaatgagctggaagaaaccccaaatggagttgctgtagtg 306  
Db 336 CTACCTGCGATTTGGAGTGTGTATCAAAAGATAAAATCGAATGTTTCAAAATCAGCAGAT 395  
QY 307 atcaaaagcgtattatggtgcaaacacttcaagctcttaatttctgcataatacaactgt 366  
Db 396 ATCAAAATCATGGCTTTTAATGACATTACGTGGGATACATTCATTCGGAATTTTATT 455  
QY 367 attcacagagataaaaaacctgaaatattcttaataactaagaagaagaataatcaagatt 426  
Db 456 TTACATCGTGAATTCNAACCAAAATTAATTTATTGGCCACCGGATGACAAATTGAAAAATA 515  
QY 427 tctgacttogggtttgacaaaattctg---attccaggagagatgctcacaccgattatgta 483  
Db 516 GCGGAATTTGGTCTTGCACGAGCTTTGGTAAATCCTAATGAAGATTATCATCTAATGTT 575  
QY 484 gctacgagatggtaccagagctcctgaactcttctgtggagagatactcagttatgcttca 543  
Db 576 GTCACCTAGATGTTATAGAGCCCTGAATTAATTTGCTGCTGACATTACACTGGAGCA 635  
QY 544 gtcgatatatgggtattggtgtgttttgcagagctcctgacagccagccactgtg 603  
Db 636 GTTGATATCTGTCATAGGTATAATTTGCTGAATTAATGCTGCAATACCTTATTTG 695  
QY 604 cctggaaaaatcagatggtggaccactttatctgataatcagaaacacaggaataaa 661  
Db 696 CCAGGTAAGATGAGCTGTGATCAATAGATGTTACATTTAGACTTATGGGACACCAA 753

RESULT 11  
US-08-874-347-1  
; Sequence 1, Application US/08874347  
; Patent No. 5863741  
; GENERAL INFORMATION:  
; APPLICANT: Limper, Andrew H.  
; APPLICANT: Leof, Edward B.  
; APPLICANT: Thomas, Charles F.  
; APPLICANT: Gustafson, Michael P.  
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
; TITLE OF INVENTION: CARINII  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C., P.A.  
; STREET: 60 South Sixth Street, Suite 3300  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/874,347  
; FILING DATE: 13-JUN-1997

us-09-671-050-11.rni

Fri May 3 11:02:08 2002

```
;
;
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 857...893
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 942...1096
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1147...1194
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1242...1254
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1311...1706
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1751...1891
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1944...2048
; OTHER INFORMATION:
;
; US-08-874-347-1
;
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 857...893
; OTHER INFORMATION: (A
; LOCATION: 942...1096
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1147...1194
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1242...1254
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1311...1706
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1751...1891
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1944...2048
; OTHER INFORMATION:
;
; US-09-093-522-1
;
; Query Match 10.9%; Score 103; DB 3; Length 2747;
; Best Local Similarity 57.4%; Pred. No. 1.1e-19;
; Matches 206; Conservative 0; Mismatches 150; Indels 3; Gaps 1;
;
; QY 293 ttgctgatgagtgatcaaaagcgcttatatggtgcaaacacacttcaagctcttaattctgtc 352
; DB 1351 TTGGAGCTGAATGATCAAAAATTATGTCTCAACTTGTATCAGGTCTTAATATTGTC 1410
; QY 353 atatacaactgtattccacagagatataaaacctgaaaaatattcttaataactaagcaag 412
; DB 1411 ATTCTCATCGTATTTCTTCATCGTGACTTGAAGCCCCAAAATCTTTATTGATCGAGAAG 1470
; QY 413 gaataacaagattgtgacttcggtttgtcac---aaattctgaltccaggagatgcct 469
; DB 1471 GAATCTTAACTTGCCTGATTTCCGGCTTGTCTCGGCATTTGGTGTCTCTTACGTGTT 1530
; QY 470 acaccgattatgactacagatggtaccagctcctgaactctgttgagagatactc 529
; DB 1531 ATACTCATGAGTGTGTTACACTTTGGTATCGGTCCAGCAAGTCTTTTAGTGGTCGAC 1590
; QY 530 agtatggtcttcacgtcatatatgtggctattggtgtgttttgcagagctcctgcag 589
; DB 1591 AATATGCACAGCACTTCATATTTGGAGCATTTGGATGTATTTTGTGTAATGGCTACAA 1650
; QY 590 gccagccactgtgcttggaataatcagatggaacacacttctgataatcagaaca 648
; DB 1651 AAAAACCGTATTTCACGCGGATTCTGAATTTGATCAATAATATTCAGAAATATTAGTCA 1709
;
; RESULT 12
; US-09-093-522-1
; Sequence 1, Application US/09093522
```

```
;
;
; Query Match 10.9%; Score 103; DB 3; Length 2747;
; Best Local Similarity 57.4%; Pred. No. 1.1e-19;
; Matches 206; Conservative 0; Mismatches 150; Indels 3; Gaps 1;
;
; QY 293 ttgctgatgagtgatcaaaagcgcttatatggtgcaaacacacttcaagctcttaattctgtc 352
; DB 1351 TTGGAGCTGAATGATCAAAAATTATGTCTCAACTTGTATCAGGTCTTAATATTGTC 1410
; QY 353 atatacaactgtattccacagagatataaaacctgaaaaatattcttaataactaagcaag 412
; DB 1411 ATTCTCATCGTATTTCTTCATCGTGACTTGAAGCCCCAAAATCTTTATTGATCGAGAAG 1470
; QY 413 gaataacaagattgtgacttcggtttgtcac---aaattctgaltccaggagatgcct 469
; DB 1471 GAATCTTAACTTGCCTGATTTCCGGCTTGTCTCGGCATTTGGTGTCTCTTACGTGTT 1530
; QY 470 acaccgattatgactacagatggtaccagctcctgaactctgttgagagatactc 529
; DB 1531 ATACTCATGAGTGTGTTACACTTTGGTATCGGTCCAGCAAGTCTTTTAGTGGTCGAC 1590
; QY 530 agtatggtcttcacgtcatatatgtggctattggtgtgttttgcagagctcctgcag 589
; DB 1591 AATATGCACAGCACTTCATATTTGGAGCATTTGGATGTATTTTGTGTAATGGCTACAA 1650
; QY 590 gccagccactgtgcttggaataatcagatggaacacacttctgataatcagaaca 648
; DB 1651 AAAAACCGTATTTCACGCGGATTCTGAATTTGATCAATAATATTCAGAAATATTAGTCA 1709
;
; RESULT 12
; US-09-093-522-1
; Sequence 1, Application US/09093522
```

Db 1351 TTGGAGCTGAAATGATCAAAAAATTTATGTCTCAACTTGTATCAGGTGTTAAATATTGTC 1410  
 QY 353 atatacaactgtatcacagagatataaaacccgtaaaatatttaataactaagcaag 412  
 Db 1411 ATTCATCATGTTCTTCATCGTGACCTTGAAGCCCCAAAATCTTCTTATGATCGAAG 1470  
 QY 413 gaataaagaattgtgacttcgggtttgcac---aaattctgattccagagatgcct 469  
 Db 1471 GAAATCTTAAACTTCTGCTGATTTCCGGCTTGTCTCGGCAATTTGGTGTCTTACGTGGTT 1530  
 QY 470 acacgattatgtacagagatgtaccagagctcctgaactcttctgtggagatactc 529  
 Db 1531 ATACTCATGAGTGTGTACACTTGTGTATGCTGTCGCAAGTCTTTTAGTGGTGGAC 1590  
 QY 530 agtatggttcctagtcgatataatggcctattgtgtgtttttgcagagctcctgaacag 589  
 Db 1591 ATATGCAACAGCACTTGATATTGGAGCATTTGATGATTTTTCGCTGAATGCTACAA 1650  
 QY 590 gccagccactgtggcctggaaaacagatgtgacccaactttatctgataatcagaaca 648  
 Db 1651 AAAAACCGTTATTTCAGCGCATTCGAAATTTGATGAAATATTCAGAATATTTAGTGCA 1709

RESULT 13  
 US-09-347-801-15  
 : Sequence 15, Application US/09347801  
 : Patent No. 6262345  
 : GENERAL INFORMATION:  
 : APPLICANT: Allen, Steve  
 : APPLICANT: Lee, Jian Ming  
 : TITLE OF INVENTION: Plant Protein Kinases  
 : FILE REFERENCE: BB-1171  
 : CURRENT APPLICATION NUMBER: US/09/347,801  
 : EARLIER FILING DATE: 1999-07-02  
 : EARLIER APPLICATION NUMBER: 60/092,438  
 : NUMBER OF SEQ ID NOS: 23  
 : SOFTWARE: Microsoft Office 97  
 : SEQ ID NO 15  
 : LENGTH: 1673  
 : TYPE: DNA  
 : ORGANISM: Triticum aestivum  
 : FEATURE:  
 : NAME/KEY: unsure  
 : LOCATION: (1349)  
 US-09-347-801-15

Query Match 10.4%; Score 98; DB 4; Length 1673;  
 Best Local Similarity 60.4%; Pred. No. 2.3e-18;  
 Matches 180; Conservative 0; Mismatches 115; Indels 3; Gaps 1;  
 QY 370 cacagagataaaacccgtaaaatattctaataactaagcaa---ggaataatcaagatt 426  
 Db 599 cacaggaatggaacacacacaaatgtttgtgtatccttaacccatcaagtaagatc 658  
 QY 427 tgtacttcgggtttgcacaaatctctgattccagagatgcctacacogattgttagct 486  
 Db 659 tgtgactttggaagtgcacaaagtcttggtacctggtgacccaacatagcatatgct 718  
 QY 487 acgagatgtaccagagctcctgaactcttctgtggagatactcagttggtcttcagtc 546  
 Db 719 tctcgtactactgtctctgctgactatattgtgcaactgaatatacaactcaata 778  
 QY 547 gatattggcctattgtgtgtttttgcagagctcctgacagccagcaactgtggcct 606  
 Db 779 gacattgttcagctggatgtgttctgcaagacttcttctgtcagcctctgtttcca 838  
 QY 607 gaaaatcagatgtgacccaactttatctgataatcagaacacttagaataatcc 664  
 Db 839 ggagagactcgggtgatcagctagtggagattatcatacaggttcttctgtgactcacaacc 896

RESULT 14  
 US-08-950-449A-18  
 : Sequence 18, Application US/08950449A  
 : Patent No. 5955366  
 : GENERAL INFORMATION:  
 : APPLICANT: Lee, John C.  
 : APPLICANT: Adams, Jerry L.  
 : APPLICANT: Gallagher, Timothy F.  
 : APPLICANT: Green, David W.  
 : APPLICANT: Heys, J. Richard  
 : APPLICANT: McDonnell, Peter  
 : APPLICANT: McNulty, Dean E.  
 : APPLICANT: Strickler, James E.  
 : APPLICANT: Young, Peter R.  
 : TITLE OF INVENTION: Drug Binding Protein  
 : NUMBER OF SEQUENCES: 19  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: SmithKline Beecham Corporation  
 : STREET: Corporate Intellectual Property/P.O. Box  
 : STREET: 1539  
 : CITY: King of Prussia  
 : STATE: PA  
 : COUNTRY: USA  
 : ZIP: 19406-0939  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/950,449A  
 : FILING DATE:  
 : CLASSIFICATION: 435  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: 08/123,175  
 : FILING DATE: 17-SEP-1993  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Hecht, Elizabeth J.  
 : REGISTRATION NUMBER: 8-41,824  
 : REFERENCE/DOCKET NUMBER: P50195-1D1  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (610) 270-5009  
 : TELEFAX: (610) 270-5090  
 : INFORMATION FOR SEQ ID NO: 18:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1100 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA  
 : HYPOTHETICAL: NO  
 : ANTI-SENSE: NO  
 US-08-950-449A-18

Query Match 10.3%; Score 97.8; DB 2; Length 1100;  
 Best Local Similarity 46.5%; Pred. No. 2.1e-18;  
 Matches 404; Conservative 0; Mismatches 447; Indels 18; Gaps 2;  
 QY 37 99gttcttatgggtgtgtattcaaatgcagaaacaaacacctctgcacaaagtagtaactgtt 96  
 Db 108 GCGCCCTATGGCTCTCTGTGTCTGCTTTTGACACAAAAACGGGTTACGTGTGGCAGTG 167  
 QY 97 aaaaattgttggaatcctgaagatgacctgtgtttaagaaaatagcactaagagaaata 156  
 Db 168 AAGAAGCTCTCCAGACCAATTTTCATGTCATTCATTCGCAAAAGAACCTACAGAACTG 227  
 QY 157 cgtatgtgaagcaattaaaacatccaaatcttgtgaacctcatcgaggttcagaga 216  
 Db 228 CGGTACTTAACATATGAAACATGAAATGTGATTGTCTGTGTGACCTTTTACACCT 287  
 QY 217 aaaagaaaatgattagtttttgatactgtgatcatacacttttaataagctgaa 276

Db 288 GCAAGGTCCTGGAGCAATTCATGATGTGTATCTGGTGACCCATCTCATGGGGCAGAT 347  
Qy 277 aaaaacccaatg-----gattgctgagtgaglgatcaaaaagcgtattatgg 324  
Db 348 CTGAACAACATTGTGAANTGTCAGAACTTACAGATGACCATGTTCACTTATCTAC 407  
Qy 325 caaacacttaagctcttaattctctgtatatacatataactgtattatccacagagataaaa 384  
Db 408 CAATTTCTCCGAGGCTAAAGTATATACATTCAGCTGCATTAATTCACAGGACCTTAAA 467  
Qy 385 cctgaaataattctataactaagcaaggaataatcaagatttggactcgggtttgca 444  
Db 468 CCTAGTATCTAGCTGTGAATCAAGACTGTGAGCTGAAGATTCTGGATTTTGGACTGC- 526  
Qy 445 caaatctgattccagagagatgctcctacacgattatgtagctacagatgtaccagct 504  
Db 527 -----TCGGCACACAGATGAATGAAATGACAGGCTAGCTGCCACTAGGTGTACAGGCT 581  
Qy 505 cctgaactcttctgggagagatactcagtgatgttcttcagtcgatatatggctattgt 564  
Db 582 CCTGAGATCATGCTGAACCTGGATGCATTACAACCCAGACAGATTGATATTGTCAGTGG 641  
Qy 565 tgtgtttttgagagctcctgacagggccagccactgtggccctggaaatcagatggag 624  
Db 642 TGCATAATGGCCGAGCTGTTGACTGGGAAGAACATGTTCTCTGGTACAGACCATATTAC 701  
Qy 625 caactttatctgataatcagaacactaggaataatcccaagacatcaatcaatcttt 684  
Db 702 CAGCTTCAGCAGATTATGCTGCTGACAGGACACCCCGCTCTATCTCATTAACAGGATG 761  
Qy 685 aagaatgaagggttttccatggcatcagatataactgagccagaagacatgaaactctt 744  
Db 762 CCAAGCATGAGGCAAGAACTATATTCACTCTTTGACTCTGATGCCCAAGATCACTTT 821  
Qy 745 gaggaagaattctcagatgttcatcctgtggtctggaacttcaatgaagggtgtctgaag 804  
Db 822 GCGAATGTTATTATTTGGTGCCAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881  
Qy 805 atgaatccagatgacagatgaactgttctccaaactcctgagagctcctactttgattct 864  
Db 882 TTGACTCAGATAGAGNATTACAGCGGCCCAAGCCCTTGACATGCCTTACTTTGCTCAG 941  
Qy 865 tttaagagggcccaatataaagaaagc 893  
Db 942 TACCAGGATCCTGATGATGAACCAAGTGC 970

RESULT 15  
US-08-950-449A-11  
; Sequence 11, Application US/08950449A  
; Patent No. 5953366  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John C.  
; APPLICANT: Adams, Jerry L.  
; APPLICANT: Gallagher, Timothy F.  
; APPLICANT: Green, David W.  
; APPLICANT: Heys, J. Richard  
; APPLICANT: McDonnell, Peter  
; APPLICANT: McNulty, Dean E.  
; APPLICANT: Strickler, James E.  
; APPLICANT: Young, Peter R.  
; TITLE OF INVENTION: Drug Binding Protein  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: Corporate Intellectual Property/P.O. Box  
; STREET: 1539  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/950,449A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/123,175  
; FILING DATE: 17-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hecht, Elizabeth J.  
; REGISTRATION NUMBER: P-41,824  
; REFERENCE/DOCKET NUMBER: P50195-1D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 270-5009  
; TELEFAX: (610) 270-5090  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3774 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; CELL TYPE: Monocyte  
; NAME/KEY: CDS  
; LOCATION: 379..1461  
; US-08-950-449A-11

Query Match 10.3%; Score 97.8; DB 2: Length 3774;  
Best Local Similarity 46.5%; Pred No. 3.7e-18;  
Matches 404; Conservative 0; Mismatches 447; Indels 18; Gaps 2;  
Qy 37 gggctcttatgggggttgattcaaatgcagaaacaaacccctctggacaagtagtagctgtt 96  
Db 459 GGGCGCTATGGCTGTGTGTCTGCTTTTGACACAAAAACGGGGTTCGTGTGGCAGTG 518  
Qy 97 aaaaaattgtggaatctgaagatgatacctctgttgaagaaaaatagcactaagagaata 156  
Db 519 AAGAGCTCTCCAGACCAATTCAGTCCATCATTCATGCGAAAAAGAACTACAGAGAACTG 578  
Qy 157 cgtatgtgaagcaattaaacatccaaactcttgaacccatcagaggtgttcaggaga 216  
Db 579 CGGTTACTTAAACATATGAACATGAAATGTGATTGGTCTGTGGACCGTTTTTACACCT 638  
Qy 217 aaaaggaataatgcatttagtttttgaataactgtgatcatcacacttttaaatgagctgaa 276  
Db 639 GCAAGGTCCTCGAGGATTCAGTATGATGTATCTGTGGTGACCCATCTCATGGGGCAGAT 698  
Qy 277 agaaacccaatg-----gagttgctgagtgagtgatcaaaagcgtattatgg 324  
Db 699 CTGAACAACATTGTGAAATGTCAGAGCTTACAGATGACCATGTTCACTGTTCTTATCTAC 758  
Qy 325 caaacacttaagctcttaattctgtcatatacataaactgtattcacagagataaaa 384  
Db 759 CAATTTCTCCGAGGCTAAAGTATATATATATTCAGCTGACATAATTCACAGGGACCTAAA 818  
Qy 385 cctgaaataattctataactaagcaaggaataatcaagatttggactcgggtttgca 444  
Db 819 CCTAGTAACTAGCTGTGAATGAAGACTGTGAGCTGGAAGATTCGTGGATTTTGGACTGCG- 877  
Qy 445 caaatctgattccagagagatgctcactacacgattatgtagctacagatgtgtaccagct 504  
Db 878 -----TCGGCACACAGATGAATGAAATGACAGGCTAGCTGGCCACTAGGTGTACAGGCT 932  
Qy 505 cctgaactcttctgggagagataactcagtatggttcttcagtcgatatatggctattgtt 564

Db 933 CCTGAGATCATGCTGAAGTGGATGATTAACACAGACAGTTGATATTTGGTCAGTGGGA 992  
 QY 565 tgtgttttttcagagctctctgacagccagccactgtgacctggaataatcagatgtggac 624  
 Db 993 TGCATATGGCCGAGCTCTTGACTGGAGAACATTTCTCTGGTACAGACATATTAAAC 1052  
 QY 625 caactttatctgataatcagaacactaggaataattatccccagacatcaatcaatcttt 684  
 Db 1053 CAGCTTCAGCAGATTATGCGTCTGACAGGACACCCCCCGCTTATCTCATTAACAGGATG 1112  
 QY 685 aaagttaacgggtttttccatggcatcagtataccctgagccagaagacatggaactctt 744  
 Db 1113 CCAAGCCATGAGCGCAAGAACTATATTCAGTCTTTGACTCAGATGCCGGAAGATGAACTTT 1172  
 QY 745 gaggaataagttctcagatgttcacctgtgctctgaaacttcacatgaagggtgtctgaaag 804  
 Db 1173 GCGAATGTATTTATTTGGTGCCAATCCCTGGCTGTGACTTGTGGAGAAAGATGCTTGTA 1232  
 QY 805 atgaatccagatgacagattaaacctgttccaaactcctggagagctcctactttgattct 864  
 Db 1233 TTGGACTCAGATAGAGAATTACAGGGCCCAAGCCCTTGCACATGCCCTACTTTGCTCAG 1292  
 QY 865 ttccaagagggcccaattaaagaaaaagc 893  
 Db 1293 TACCAGATCCTGATGATGAACACAGTGGC 1321

Search completed: May 2, 2002, 17:10:02  
 Job time: 9674 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 18:43:31 ; Search time 170.24 Seconds  
(without alignments)  
5242.455 Million cell updates/sec

Title: US-09-671-050-9  
Perfect score: 1041  
Sequence: 1 atggaaaagtatgaaaaatt.....aggtacttcgcgtcaaaagt 1041

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size : 20

Total number of hits satisfying chosen parameters: 781

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :				N_Geneseq_1101.*			
1:	/SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*	1:	/SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*	1:	/SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*	1:	/SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*	2:	/SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*	2:	/SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*	2:	/SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*	3:	/SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*	3:	/SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*	3:	/SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*	4:	/SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*	4:	/SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*	4:	/SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*	5:	/SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*	5:	/SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*	5:	/SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*	6:	/SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*	6:	/SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*	6:	/SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*	7:	/SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*	7:	/SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*	7:	/SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*	8:	/SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*	8:	/SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*	8:	/SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*	9:	/SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*	9:	/SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*	9:	/SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*	10:	/SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*	10:	/SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*	10:	/SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*	11:	/SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*	11:	/SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*	11:	/SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*	12:	/SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*	12:	/SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*	12:	/SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*	13:	/SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*	13:	/SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*	13:	/SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*	14:	/SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*	14:	/SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*	14:	/SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*	15:	/SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*	15:	/SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*	15:	/SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*	16:	/SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*	16:	/SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*	16:	/SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*	17:	/SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*	17:	/SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*	17:	/SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*	18:	/SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*	18:	/SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*	18:	/SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*	19:	/SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*	19:	/SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*	19:	/SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*	20:	/SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*	20:	/SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*	20:	/SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*	21:	/SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*	21:	/SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*	21:	/SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*	22:	/SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*	22:	/SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*	22:	/SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	1041	22	Human kinase cDNA
2	1025	98.5	1068	22	Human kinase cDNA
3	652	62.6	945	22	Human kinase cDNA
4	652	62.6	972	22	Human kinase cDNA
5	572	54.9	1819	22	Human kinase cDNA
6	454	43.6	561	22	Human kinase cDNA
7	454	43.6	594	22	Human kinase cDNA
8	454	43.6	1083	22	Polynucleotide seq
9	368	35.4	911	22	Novel protein kina
10	40	3.8	12729	22	Human neuroblastom
11	33	3.2	105	21	Human secreted pro

12	33	3.2	411	14	AAO61041	Human brain Expres
13	33	3.2	2286	22	AAF26596	DNA encoding human
14	33	3.2	5600	22	AAI60415	Human polynucleoti
15	33	3.2	6027	22	AAI58629	Human polynucleoti
16	31	3.0	32191	22	AAI64117	Human bladder rela
17	30	2.9	300	20	AAI58629	Human cancer cell
18	30	2.9	752	20	AAI58629	Human validated ca
19	30	2.9	1345	21	AAI58629	Human secreted pro
20	30	2.9	1718	22	AAI58629	Human cervical can
21	30	2.9	1762	22	AAI58629	Mammalian toxicolo
22	30	2.9	1804	22	AAI58629	Human cDNA sequenc
23	30	2.9	1916	22	AAI58629	Human kidney relat
24	29	2.8	9439	22	AAI58629	Human kidney relat
25	29	2.8	237326	19	AAI58629	Hereditary haemoch
26	28	2.7	261	21	AAI58629	Human secreted pro
27	28	2.7	323	21	AAI58629	Human secreted pro
28	28	2.7	531	22	AAI58629	Human cDNA clone (
29	28	2.7	631	21	AAI58629	Human cDNA clone (
30	28	2.7	805	22	AAI58629	Human GSHS related
31	28	2.7	1001	21	AAI58629	Human GSHS related
32	28	2.7	1001	21	AAI58629	Nucleotide sequenc
33	28	2.7	1469	22	AAI58629	Human transmembran
34	28	2.7	1562	21	AAI58629	Human EST-derived
35	28	2.7	1624	22	AAI58629	Nucleotide sequenc
36	28	2.7	1763	22	AAI58629	Nucleotide sequenc
37	28	2.7	1763	22	AAI58629	Human cDNA sequenc
38	28	2.7	1791	22	AAI58629	Human cDNA sequenc
39	28	2.7	1792	22	AAI58629	Human secreted pro
40	28	2.7	1850	21	AAI58629	Human ORFX ORF2491
41	28	2.7	1902	21	AAI58629	Human cDNA sequenc
42	28	2.7	2252	22	AAI58629	Human chemokine al
43	28	2.7	2309	20	AAI58629	Human CXCL family c
44	28	2.7	2575	21	AAI58629	Human cDNA sequenc
45	28	2.7	2797	22	AAI58629	Human CXCL family c
46	28	2.7	2797	22	AAI58629	Human cDNA sequenc
47	28	2.7	2845	22	AAI58629	Human low adenosin
48	28	2.7	3585	21	AAI58629	Human adenosine re
49	28	2.7	3585	21	AAI58629	Nitric oxide synth
50	28	2.7	3586	22	AAI58629	Human cDNA sequenc
51	28	2.7	3917	22	AAI58629	Human kidney relat
52	28	2.7	9456	22	AAI58629	Human colorectal c
53	28	2.7	16552	22	AAI58629	Human colorectal c
54	28	2.7	16552	22	AAI58629	Human neuroblastom
55	28	2.7	23241	22	AAI58629	Human neuroblastom
56	28	2.7	23241	22	AAI58629	Human LOBO homolog
57	28	2.7	30417	21	AAI58629	Human adenosine re
58	28	2.7	30417	21	AAI58629	Human LOBO homolog
59	28	2.7	49959	20	AAI58629	Nucleotide sequenc
60	28	2.7	49959	20	AAI58629	Human neuroblastom
61	28	2.7	50000	21	AAI58629	DNA sequence of th
62	28	2.7	51474	22	AAI58629	NC-contig derived
63	28	2.7	54548	21	AAI58629	NC-contig derived
64	28	2.7	80240	20	AAI58629	Human PACAP genom
65	28	2.7	80595	20	AAI58629	Human TBC-1 partia
66	28	2.7	84607	20	AAI58629	Human kinesin-like
67	28	2.7	99960	21	AAI58629	Human KCMQ5 (KCN6q
68	28	2.7	121162	21	AAI58629	Human hyporetin r
69	28	2.7	125910	21	AAI58629	Human PG-3 gene.
70	28	2.7	168575	22	AAI58629	Human chromosome 1
71	28	2.7	240825	22	AAI58629	Human schizopreni
72	28	2.7	319608	21	AAI58629	Human secreted pro
73	28	2.7	319608	22	AAI58629	Human secreted pro
74	28	2.7	319608	22	AAI58629	Human secreted pro
75	27	2.6	182	21	AAI58629	Human secreted pro
76	27	2.6	182	21	AAI58629	Human secreted pro
77	27	2.6	183	21	AAI58629	Human secreted pro
78	27	2.6	215	21	AAI58629	Human brain Expres
79	27	2.6	278	14	AAI58629	Human MN gene Alu-
80	27	2.6	294	21	AAI58629	MN genomic DNA seq
81	27	2.6	294	21	AAI58629	Human colon cancer
82	27	2.6	366	22	AAI58629	Human cDNA clone (
83	27	2.6	529	22	AAI58629	Human cDNA clone (
84	27	2.6	536	22	AAI58629	Human cDNA clone (

85	27	2.6	562	22	AAH10364	Human cDNA clone (	158	27	2.6	141589	21	AAF21152	Human low adenosin
86	27	2.6	563	22	AAH10427	Human cDNA clone (	159	27	2.6	141589	21	AAA35005	Human adenosine re
87	27	2.6	566	22	AAH10170	Human cDNA clone (	160	27	2.6	141589	21	AAA35030	Human adenosine re
88	27	2.6	588	22	AAH09479	Human cDNA clone (	161	27	2.6	146981	21	AAF21442	Human ELAM-1 polyn
89	27	2.6	616	20	AAH03004	Human IL-1ra BAC c	162	27	2.6	161425	22	AAH02340	Human AKAP10 gene
90	27	2.6	698	22	AAH03391	Human cDNA clone (	163	27	2.6	162025	22	AAH02339	Human AKAP10 gene
91	27	2.6	700	22	AAH92306	Human inflammatory	164	27	2.6	209273	21	AAF21437	Human factor relat
92	27	2.6	778	21	AAH15909	Human prostate can	165	26	2.5	176	22	AAI63308	Human kidney relat
93	27	2.6	1162	21	AAH74355	Human secreted pro	166	26	2.5	317	21	AAH29902	Human secreted pro
94	27	2.6	1174	21	AAH52485	Human MN gene intr	167	26	2.5	397	21	AAH66996	Human secreted pro
95	27	2.6	1174	21	AAH16566	Human MN gene intr	168	26	2.5	837	20	AAH37525	Human secreted pro
96	27	2.6	1255	21	AAH5487	Human secreted pro	169	26	2.5	954	21	AAH66442	Human secreted pro
97	27	2.6	1281	21	AAH5931	Membrane-bound pro	170	26	2.5	1001	21	AAH57767	Human secreted pro
98	27	2.6	1281	22	AAH5931	Human cDNA sequenc	171	26	2.5	1067	21	AAH59697	Arachidonic acid m
99	27	2.6	1281	22	AAH5931	Human PRO1198 (UNQ	172	26	2.5	1643	22	AAH18713	Human secreted pro
100	27	2.6	1297	21	AAH44238	Human prostate can	173	26	2.5	1784	22	AAH87818	Human cDNA sequenc
101	27	2.6	1300	22	AAH64835	Human secreted pro	174	26	2.5	1968	22	AAH18031	Human haemopoietin
102	27	2.6	1301	22	AAH64834	Human secreted pro	175	26	2.5	2580	22	AAH17178	Human cDNA sequenc
103	27	2.6	1320	22	AAH15250	Human secreted pro	176	26	2.5	3292	21	AAH77356	Human ORFX ORF2911
104	27	2.6	1334	21	AAH1655	Human cDNA sequenc	177	26	2.5	3906	22	AAH73707	Nucleotide sequenc
105	27	2.6	1361	21	AAH93821	N. meningitidis pa	178	26	2.5	5004	22	AAH27885	Human low adenosin
106	27	2.6	1401	21	AAH52495	Human cDNA encodin	179	26	2.5	5683	21	AAH21125	Human adenosine re
107	27	2.6	1401	21	AAH16576	Human MN gene frag	180	26	2.5	59960	21	AAH250905	Human TBC-1 partia
108	27	2.6	1447	22	AAH159422	Human polynucleoti	181	26	2.5	99960	21	AAH250905	Human ABC1 genom
109	27	2.6	1448	22	AAH17022	Human cDNA sequenc	182	26	2.5	162025	22	AAH02339	Hereditary haemoch
110	27	2.6	1609	19	AAH17088	Homo sapiens tubul	183	26	2.5	183999	22	AAH02339	Hereditary haemoch
111	27	2.6	1613	19	AAH59193	Human cDNA sequenc	184	26	2.5	235033	19	AAH57926	Human secreted pro
112	27	2.6	1879	21	AAH75943	Human SR-BI recept	185	26	2.5	237326	19	AAH57926	Human secreted pro
113	27	2.6	1901	20	AAH201024	Human ORFX ORF1498	186	25	2.4	251	21	AAH23479	Human lung tumour
114	27	2.6	2084	22	AAH17430	Partial human PGI	187	25	2.4	261	21	AAH51124	Human secreted pro
115	27	2.6	2177	22	AAH17542	Human cDNA sequenc	188	25	2.4	307	21	AAH51125	Human MGST-II rela
116	27	2.6	2234	20	AAH200451	Human cDNA sequenc	189	25	2.4	341	21	AAH51125	Secreted protein E
117	27	2.6	2442	22	AAH45208	Human secreted pro	190	25	2.4	341	21	AAH51125	Human cDNA sequenc
118	27	2.6	3598	22	AAH18642	Human cytochrome C	191	25	2.4	341	22	AAH51125	Human kidney relat
119	27	2.6	4153	22	AAH18395	Human cDNA sequenc	192	25	2.4	346	21	AAH51125	Human kidney relat
120	27	2.6	4823	16	AAH03943	Human cDNA sequenc	193	25	2.4	884	21	AAH51125	Human IL-1ra BAC c
121	27	2.6	4823	16	AAH04051	Human thrombopoiet	194	25	2.4	1289	20	AAH51125	Human cDNA sequenc
122	27	2.6	4823	17	AAH37384	Human thrombopoiet	195	25	2.4	1983	22	AAH18571	Human secreted pro
123	27	2.6	4823	21	AAH51993	Human thrombopoiet	196	25	2.4	3369	22	AAH18571	Human secreted pro
124	27	2.6	4823	21	AAH37785	Human thrombopoiet	197	25	2.4	3370	22	AAH18571	Human secreted pro
125	27	2.6	5117	22	AAH16663	Human cDNA sequenc	198	25	2.4	4696	20	AAH03041	Human cDNA sequenc
126	27	2.6	6268	22	AAH01148	Interferon induced	199	25	2.4	5488	22	AAH18214	Human cDNA sequenc
127	27	2.6	7676	19	AAH62906	Human galactokinase	200	25	2.4	12312	22	AAH18214	Human neuroblastom
128	27	2.6	7849	16	AAH094109	hML genomic DNA.	201	25	2.4	18647	21	AAH18214	Human low adenosin
129	27	2.6	10475	9	AAH80616	Genomic clone enco	202	25	2.4	18648	21	AAH18214	Human low adenosin
130	27	2.6	10475	9	AAH80616	Sequence of human	203	25	2.4	25235	22	AAH18214	Human adenosine re
131	27	2.6	10884	21	AAH66549	Human kinesin-like	204	25	2.4	50000	21	AAH18214	Polymorphic repeat
132	27	2.6	10897	17	AAH09187	Mutu putative onco	205	25	2.4	11950	20	AAH18214	Human yes1 gene.
133	27	2.6	10898	21	AAH52462	Human MN gene. Ho	206	24	2.3	129	21	AAH18214	Human secreted pro
134	27	2.6	10898	21	AAH16543	Human MN complete	207	24	2.3	210	21	AAH18214	Human secreted pro
135	27	2.6	11992	22	AAH16543	Human kidney relat	208	24	2.3	305	21	AAH18214	Human secreted pro
136	27	2.6	12683	22	AAH16543	Human kidney relat	209	24	2.3	411	20	AAH18214	EST Clone 0372. H
137	27	2.6	12744	22	AAH16543	Nucleotide sequenc	210	24	2.3	466	21	AAH18214	Human secreted pro
138	27	2.6	14784	21	AAH64141	Human breast or ov	211	24	2.3	484	22	AAH18214	Human cDNA clone (
139	27	2.6	21441	22	AAH62667	Human breast or ov	212	24	2.3	534	22	AAH09869	Human cDNA clone (
140	27	2.6	32174	22	AAH62667	Human genomic DNA	213	24	2.3	550	22	AAH13330	Human cDNA clone (
141	27	2.6	32174	22	AAH62667	Human neuroblastom	214	24	2.3	777	22	AAH04994	cDNA encoding a hu
142	27	2.6	34488	22	AAH62667	Human PKD1 gene.	215	24	2.3	812	20	AAH04994	Human cDNA sequenc
143	27	2.6	37808	20	AAH62667	Human polycystic k	216	24	2.3	1046	22	AAH18696	Human low adenosin
144	27	2.6	37808	20	AAH62667	Human PKD1 locus b	217	24	2.3	1520	21	AAH18696	Human adenosine re
145	27	2.6	53526	19	AAH62667	PGI genomic coding	218	24	2.3	1520	21	AAH18696	Human secreted pro
146	27	2.6	53577	17	AAH18551	Wild type PGI codi	219	24	2.3	1572	21	AAH18696	Human colon cancer
147	27	2.6	53577	17	AAH18551	Human thioredoxin	220	24	2.3	1586	21	AAH18696	NADH dehydrogenase
148	27	2.6	56516	20	AAH00870	Human genomic DNA	221	24	2.3	1669	22	AAH18696	First intron promo
149	27	2.6	56520	20	AAH00870	Human FN gene. Ho	222	24	2.3	1827	17	AAH18696	Human kidney relat
150	27	2.6	66566	21	AAH53450	Human PCTA-1 genom	223	24	2.3	1827	17	AAH18696	Human cDNA sequenc
151	27	2.6	66566	22	AAH53450	Human kinesin-like	224	24	2.3	1961	22	AAH18696	Human cDNA sequenc
152	27	2.6	78925	21	AAH07380	Human histone deac	225	24	2.3	2088	22	AAH18696	Human low adenosin
153	27	2.6	106746	21	AAH10225	Human adenosine re	226	24	2.3	2277	21	AAH18696	Human cDNA sequenc
154	27	2.6	122186	22	AAH89560	Human ELAM-1 polyn	227	24	2.3	2285	22	AAH18696	Human cDNA sequenc
155	27	2.6	138169	21	AAH34791	Human low adenosin	228	24	2.3	2492	22	AAH18696	Human kidney relat
156	27	2.6	141589	21	AAH20913		229	24	2.3	2716	22	AAH18696	
157	27	2.6	141589	21	AAH21127		230	24	2.3				

c 231	24	2.3	2716	22	AAI63635	Human kidney relat	c 304	23	2.2	11725	18	AAI79723	Human protein C ge
c 232	24	2.3	2945	21	AAI68095	Human secreted pro	c 305	23	2.2	11725	20	AAI32179	Human protein C nu
c 233	24	2.3	2945	21	AAI68095	Human colon cancer	c 306	23	2.2	11725	20	AAI32179	Human protein C CD
c 234	24	2.3	3073	22	AAH33162	Human cDNA sequenc	c 307	23	2.2	36221	22	AAI54051	Human death-associ
c 235	24	2.3	3871	20	AAH35082	Human cDNA sequenc	c 308	23	2.2	41684	21	AAI00624	Human purH gene ge
c 236	24	2.3	4171	22	AAH160801	Human cDNA sequenc	c 309	23	2.2	44848	21	AAI57508	Nucleotide sequenc
c 237	24	2.3	4365	21	AAH160801	Human cDNA sequenc	c 310	23	2.2	144460	21	AAI93815	Olfactory receptor
c 238	24	2.3	4803	20	AAI291172	Human vanilloid re	c 311	23	2.2	183999	22	AAI92831	Human ABC1 genomic
c 239	24	2.3	4803	20	AAI291172	Human vanilloid re	c 312	23	2.2	117	22	AAI12107	Human secreted pro
c 240	24	2.3	4803	20	AAI291172	Human vanilloid re	c 313	23	2.2	125	21	AAI14035	Human secreted pro
c 241	24	2.3	4824	21	AAI66464	Human vanilloid re	c 314	23	2.2	183	21	AAI20144	Human gene signatu
c 242	24	2.3	8559	22	AAI63361	Human kidney relat	c 315	23	2.2	189	16	AAI24873	Human secreted pro
c 243	24	2.3	8864	21	AAI63361	Human kidney relat	c 316	23	2.2	249	21	AAI21715	Human secreted pro
c 244	24	2.3	8864	21	AAI63361	Human kidney relat	c 317	23	2.2	289	21	AAI04999	Human secreted pro
c 245	24	2.3	9734	22	AAI64136	Human adiposine re	c 318	23	2.2	293	21	AAI26337	Human secreted pro
c 246	24	2.3	11617	22	AAI62937	Human lipoprotein	c 319	23	2.2	300	20	AAI13542	Human gene express
c 247	24	2.3	13204	20	AAI32189	Human genomic DNA	c 320	23	2.2	302	21	AAI02929	Human secreted pro
c 248	24	2.3	24025	17	AAI17455	Human platelet gly	c 321	23	2.2	334	21	AAI19789	Human secreted pro
c 249	24	2.3	24025	17	AAI17455	Human platelet gly	c 322	23	2.2	334	21	AAI19789	Human secreted pro
c 250	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 323	23	2.2	357	21	AAI16430	Human secreted pro
c 251	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 324	23	2.2	365	21	AAI15064	Human secreted pro
c 252	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 325	23	2.2	382	14	AAI39885	Expressed Sequence
c 253	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 326	23	2.2	382	14	AAI39885	Human brain Expres
c 254	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 327	23	2.2	495	22	AAI52927	Human secreted pro
c 255	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 328	23	2.2	527	21	AAI08525	Human secreted exp
c 256	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 329	23	2.2	527	21	AAI08525	N. meningitidis pa
c 257	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 330	23	2.2	540	21	AAI1688	Human CDNA clone (
c 258	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 331	23	2.2	559	22	AAI10229	Human CDNA clone (
c 259	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 332	23	2.2	570	22	AAI13418	Human inflammatory
c 260	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 333	23	2.2	700	22	AAI92232	Human inflammatory
c 261	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 334	23	2.2	778	22	AAI07224	Human CDNA clone (
c 262	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 335	23	2.2	789	21	AAI07224	Human secreted pro
c 263	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 336	23	2.2	843	22	AAI04635	Human cytoplasmic
c 264	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 337	23	2.2	877	19	AAI04635	Human cytoplasmic
c 265	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 338	23	2.2	1001	21	AAI57479	Arachidonic acid m
c 266	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 339	23	2.2	1098	19	AAI57479	Homo sapiens BAR1
c 267	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 340	23	2.2	1495	22	AAI08502	Human secreted pro
c 268	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 341	23	2.2	1533	21	AAI21366	Human low adenosin
c 269	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 342	23	2.2	1533	21	AAI21366	Human adenosine re
c 270	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 343	23	2.2	1774	22	AAI99647	Human protein enco
c 271	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 344	23	2.2	1784	21	AAI99647	Nucleotide sequenc
c 272	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 345	23	2.2	1816	22	AAI72758	Human prostate can
c 273	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 346	23	2.2	1848	22	AAI16421	Human CDNA sequenc
c 274	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 347	23	2.2	1880	22	AAI16035	Human CDNA sequenc
c 275	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 348	23	2.2	1964	22	AAI15328	Human CDNA sequenc
c 276	24	2.3	24026	17	AAI17512	Mutated BRCAL geno	c 349	23	2.2	2157	21	AAI74405	Human secreted pro
c 277	23	2.2	351	22	AAI11840	Human secreted pro	c 350	23	2.2	2176	22	AAI15589	Human oligodendro
c 278	23	2.2	351	22	AAI11840	Human secreted pro	c 351	23	2.2	2243	22	AAI89709	Human CDNA sequenc
c 279	23	2.2	374	14	AAI059566	Human brain Expres	c 352	23	2.2	2271	22	AAI17984	Human secreted pro
c 280	23	2.2	390	21	AAI059566	Human brain Expres	c 353	23	2.2	2325	22	AAI07578	Human CRE1 nucleot
c 281	23	2.2	407	21	AAH30958	Human secreted pro	c 354	23	2.2	2337	21	AAI29581	Human polynucleoti
c 282	23	2.2	419	21	AAH31126	Human colon cancer	c 355	23	2.2	2350	22	AAI15637	Human CDNA sequenc
c 283	23	2.2	526	16	AAQ86309	Human colon cancer	c 356	23	2.2	2548	22	AAI160131	Human CDNA sequenc
c 284	23	2.2	556	22	AAH16931	Wilson disease gen	c 357	23	2.2	3161	22	AAI17293	Human polynucleoti
c 285	23	2.2	577	22	AAH09267	Human CDNA clone (	c 358	23	2.2	3161	22	AAI17293	Human CDNA sequenc
c 286	23	2.2	688	17	AAI16856	Integrin subunit b	c 359	23	2.2	3191	21	AAI35246	Human low adenosin
c 287	23	2.2	698	17	AAI16856	Human colon cancer	c 360	23	2.2	3191	21	AAI35246	Human adenosine re
c 288	23	2.2	700	22	AAH29897	Human inflammatory	c 361	23	2.2	3303	22	AAH18287	Human cDNA sequenc
c 289	23	2.2	1086	22	AAH45109	Human inflammatory	c 362	23	2.2	3885	19	AAI25508	Human bax promoter
c 290	23	2.2	1741	22	AAH15292	Human CDNA sequenc	c 363	23	2.2	4041	22	AAH46858	Human serine/threo
c 291	23	2.2	1798	22	AAH13835	Human CDNA sequenc	c 364	23	2.2	4156	22	AAH57458	Human liver cell s
c 292	23	2.2	1871	22	AAH17838	Human secreted pro	c 365	23	2.2	4157	22	AAH57457	Human kidney relat
c 293	23	2.2	2504	22	AAH17838	Human secreted pro	c 366	23	2.2	6775	22	AAI3671	Glucocerebrosidase
c 294	23	2.2	2615	22	AAI44670	Human CDNA sequenc	c 367	23	2.2	7620	14	AAI39286	Human kidney relat
c 295	23	2.2	3065	21	AAI00128	Human NIP2b cDNA	c 368	23	2.2	11992	22	AAI63357	DNA sequence encod
c 296	23	2.2	3227	22	AAH16625	Human NIP2b cDNA	c 369	23	2.2	12151	21	AAI63357	Human genomic DNA
c 297	23	2.2	5124	21	AAI75761	Human ORF ORF1316	c 370	23	2.2	14417	22	AAI62923	Human genomic DNA
c 298	23	2.2	5124	21	AAI75761	Human ORF ORF1316	c 371	23	2.2	14426	22	AAI62923	Human genomic DNA
c 299	23	2.2	6405	22	AAI75761	Human ORF ORF1316	c 372	23	2.2	14426	22	AAI62923	Human genomic DNA
c 300	23	2.2	10080	21	AAI95933	Human KIK-L4 gene	c 373	23	2.2	14451	22	AAI62923	Human genomic DNA
c 301	23	2.2	11715	9	AAH1564	Genomic sequence e	c 374	23	2.2	14451	22	AAI62923	Human genomic DNA
c 302	23	2.2	11724	22	AAH70102	Complete genomic s	c 375	23	2.2	14483	22	AAH45311	Human HCR DNA. HO
c 303	23	2.2	11725	17	AAI32796	Human liver cell s	c 376	23	2.2	14796	19	AAI27941	Survivin gene. HO

377	22	2.1	15071	21	AAZ58411	hnpA2 genomic c	450	21	2.0	35651	22	AAF57595	ATM complete genom
378	22	2.1	16163	22	AAI62620	Human breast or ov	451	21	2.0	36221	22	AA500624	Human death-associi
379	22	2.1	17327	14	AAO44278	Serglycin - proteo	c 452	21	2.0	66685	22	AA507380	Human genomic DNA
c 380	22	2.1	18402	21	AAZ93705	CyB5RP fatty acid	c 453	21	2.0	68940	20	AA573531	Human chromosome 6
c 381	22	2.1	28871	21	AAO09888	Human genomic OCTN	c 454	20	1.9	119	21	AAZ12930	Human secreted pro
c 382	22	2.1	35465	22	AAF54723	Nucleotide sequenc	c 455	20	1.9	139	21	AAZ28706	Human secreted pro
c 383	22	2.1	72604	20	AAZ10752	Genomic sequence o	c 456	20	1.9	151	21	AAZ27177	Human secreted pro
c 384	22	2.1	72604	20	AAZ10752	Genomic sequence o	c 457	20	1.9	161	21	AAZ25657	Human secreted pro
c 385	22	2.1	80240	20	AAV83940	NC-contig derived	c 458	20	1.9	167	21	AAZ18081	Human secreted pro
c 386	22	2.1	80595	20	AAV83939	NC-contig derived	c 459	20	1.9	170	21	AAZ14545	Human secreted pro
c 387	22	2.1	110000	22	AAF84800	Nucleotide sequenc	c 460	20	1.9	174	21	AAZ14545	Human secreted pro
c 388	21	2.0	107	21	AAZ14249	Human secreted pro	c 461	20	1.9	174	21	AAZ14545	Human secreted pro
c 389	21	2.0	122	21	AAZ13954	Human secreted pro	c 462	20	1.9	182	21	AAZ08211	Human secreted exp
c 390	21	2.0	170	21	AAZ23778	Human secreted pro	c 463	20	1.9	193	21	AAZ15647	Human secreted pro
c 391	21	2.0	188	21	AAZ15303	Human secreted pro	c 464	20	1.9	198	21	AAZ27968	Human secreted pro
c 392	21	2.0	232	22	AAI62838	Human secreted pro	c 465	20	1.9	211	21	AAZ20456	Human secreted pro
c 393	21	2.0	239	21	AAZ01018	Human genomic DNA	c 466	20	1.9	212	21	AAZ14484	Human secreted pro
c 394	21	2.0	251	21	AAZ04928	Human secreted pro	c 467	20	1.9	221	21	AAZ14397	Human secreted pro
c 395	21	2.0	251	21	AAZ04928	Human secreted pro	c 467	20	1.9	252	21	AAZ32419	Human secreted pro
c 396	21	2.0	258	22	AAI20461	Probe #10394 for g	c 468	20	1.9	264	17	AAQ12395	3'-Flanking sequen
c 397	21	2.0	258	22	AAI05667	Probe #14353 used	c 469	20	1.9	265	14	AAQ39724	Expressed Sequence
c 398	21	2.0	259	21	AAI06160	Probe #6151 used t	c 470	20	1.9	265	14	AAQ59136	Human brain Expres
c 399	21	2.0	259	21	AAZ13236	Human secreted pro	c 471	20	1.9	275	22	AAH71358	Human cervical can
c 400	21	2.0	259	21	AAZ13333	Human secreted pro	c 472	20	1.9	286	21	AAH71358	Human secreted pro
c 401	21	2.0	259	21	AAZ15679	Human secreted pro	c 473	20	1.9	288	21	AAZ13443	Human secreted pro
c 402	21	2.0	259	21	AAZ15679	Human secreted pro	c 474	20	1.9	288	21	AAZ13443	Human secreted pro
c 403	21	2.0	260	21	AAZ15764	Human secreted pro	c 475	20	1.9	292	16	AAZ22881	Human secreted pro
c 404	21	2.0	311	21	AAZ15732	Human secreted pro	c 476	20	1.9	292	16	AAZ22881	Human secreted pro
c 405	21	2.0	343	21	AAZ10694	Human secreted pro	c 477	20	1.9	299	22	AAZ22926	Human secreted pro
c 406	21	2.0	369	21	AAZ15656	Human secreted pro	c 478	20	1.9	300	20	AAZ163724	Human brain Expres
c 407	21	2.0	370	14	AAQ60235	Human prostate can	c 479	20	1.9	304	20	AAZ163724	Human cervical can
c 408	21	2.0	467	22	AAZ97263	Human brain Expres	c 480	20	1.9	312	16	AAZ24725	Human gene express
c 409	21	2.0	479	22	AAI11240	Human prostate can	c 481	20	1.9	326	21	AAZ10629	Human secreted pro
c 410	21	2.0	479	22	AAI01156	Probe #1173 for 9e	c 482	20	1.9	327	22	AAH69933	Human secreted pro
c 411	21	2.0	554	22	AAH11403	Probe #1191 used t	c 483	20	1.9	330	21	AAZ69933	Human secreted pro
c 412	21	2.0	585	22	AAH110314	Probe #1147 used t	c 484	20	1.9	341	22	AAH19245	Human secreted pro
c 413	21	2.0	588	22	AAH10028	Human cDNA clone (	c 485	20	1.9	350	21	AAZ08582	Human secreted pro
c 414	21	2.0	700	22	AAH92149	Human cDNA clone (	c 486	20	1.9	373	21	AAZ17118	Human secreted pro
c 415	21	2.0	700	22	AAH92150	Human inflammatory	c 487	20	1.9	373	21	AAZ17118	Human secreted pro
c 416	21	2.0	700	22	AAH93077	Human inflammatory	c 488	20	1.9	376	21	AAH30707	Human secreted pro
c 417	21	2.0	700	22	AAH93078	Human inflammatory	c 489	20	1.9	377	20	AAV87212	Human secreted pro
c 418	21	2.0	888	21	AAZ59204	Human secreted pro	c 490	20	1.9	383	22	AAZ67478	Human secreted pro
c 419	21	2.0	1376	22	AAZ60800	Human polynucleoti	c 491	20	1.9	393	14	AAO60969	Novel human polynu
c 420	21	2.0	1419	22	AAZ21982	Human polynucleoti	c 492	20	1.9	397	22	AAZ60969	Novel human polynu
c 421	21	2.0	1545	22	AAZ160807	Human collagen gen	c 493	20	1.9	399	14	AAO61226	Human brain Expres
c 422	21	2.0	1632	22	AAH17196	Human polynucleoti	c 494	20	1.9	401	14	AAO61312	Human brain Expres
c 423	21	2.0	1634	22	AAH81056	Human cDNA sequenc	c 495	20	1.9	403	21	AAZ06696	Human secreted pro
c 424	21	2.0	1650	22	AAZ33121	Human secreted pro	c 496	20	1.9	407	21	AAH30467	Human secreted pro
c 425	21	2.0	1719	22	AAH18317	Human secreted pro	c 497	20	1.9	413	22	AAH30387	Human colon cancer
c 426	21	2.0	1768	22	AAZ23896	Human cDNA sequenc	c 498	20	1.9	423	21	AAH30387	Human colon cancer
c 427	21	2.0	1830	21	AAZ21430	Human secreted pro	c 499	20	1.9	425	21	AAH30501	Human secreted pro
c 428	21	2.0	1830	21	AAZ21430	Human secreted pro	c 500	20	1.9	425	21	AAH30501	Human secreted pro
c 429	21	2.0	1901	22	AAH35308	Human low adenosin	c 501	20	1.9	430	19	AAZ60500	Human colon cancer
c 430	21	2.0	2036	22	AAH17363	Human cDNA sequenc	c 502	20	1.9	430	19	AAZ60500	Human colon cancer
c 431	21	2.0	2085	21	AAZ34408	DNA encoding human	c 503	20	1.9	432	22	AAH08962	Human brain Expres
c 432	21	2.0	2529	18	AAZ87468	Hamster Ubiquitin/	c 504	20	1.9	445	22	AAH04811	DNA molecule encod
c 433	21	2.0	2545	22	AAH17805	Human cDNA sequenc	c 505	20	1.9	458	21	AAZ06059	Human breast tumou
c 434	21	2.0	2863	22	AAH17480	Human cDNA sequenc	c 506	20	1.9	464	21	AAZ06059	Novel human polynu
c 435	21	2.0	2918	21	AAZ73083	Human cDNA sequenc	c 507	20	1.9	479	22	AAH04811	Human cDNA clone (
c 436	21	2.0	2960	22	AAZ63590	Human cDNA sequenc	c 508	20	1.9	479	22	AAH04811	Human cDNA clone (
c 437	21	2.0	3138	22	AAZ02640	Human kidney relat	c 509	20	1.9	499	22	AAH08227	Human secreted pro
c 438	21	2.0	3193	22	AAH18573	Human kidney relat	c 510	20	1.9	499	22	AAH08227	Human colon cancer
c 439	21	2.0	3204	21	AAH73082	Human cDNA sequenc	c 511	20	1.9	523	21	AAH09770	Human cDNA clone (
c 440	21	2.0	3617	22	AAH13351	Human melanocortin	c 512	20	1.9	523	21	AAH09770	Human cDNA clone (
c 441	21	2.0	3642	22	AAZ03039	Human IL-1ra BAC c	c 513	20	1.9	545	22	AAH13490	Human secreted exp
c 442	21	2.0	3706	22	AAH18728	Human cDNA sequenc	c 514	20	1.9	546	22	AAH13490	Human cDNA clone (
c 443	21	2.0	6246	22	AAZ22408	Human cDNA sequenc	c 515	20	1.9	549	22	AAH09726	Human cDNA clone (
c 444	21	2.0	6246	22	AAZ22408	Human cDNA sequenc	c 516	20	1.9	554	22	AAH09726	Human cDNA clone (
c 445	21	2.0	6511	16	AAO95493	Nucleotide sequenc	c 517	20	1.9	556	21	AAH13313	Human cDNA clone (
c 446	21	2.0	6546	21	AAZ14133	Human cDNA sequenc	c 518	20	1.9	556	21	AAH13313	Human secreted exp
c 447	21	2.0	6546	21	AAZ14133	Human cDNA sequenc	c 519	20	1.9	561	22	AAH13244	Human cDNA clone (
c 448	21	2.0	15577	19	AAZ35616	Human adenosine re	c 520	20	1.9	565	22	AAH09959	Human cDNA clone (
c 449	21	2.0	32152	22	AAI57791	SHOX gene prelimin	c 521	20	1.9	576	22	AAH081826	Alternatively spli
c 450	21	2.0	32152	22	AAI57791	Human colorectal c	c 522	20	1.9	595	22	AAH09633	Human cDNA clone (
c 451	21	2.0	32152	22	AAI57791	Human colorectal c	c 522	20	1.9	595	22	AAH09633	Human EST-derived

us-09-671-050-9.oli20.rng

Fri May 3 11:02:10 2002

523	20	1.9	603	22	AAH03596	Human cDNA clone (	596	20	1.9	1811	22	AAH05407	Human secreted pro
524	20	1.9	614	20	AAV98921	EST clone HW591.	597	20	1.9	1821	22	AAH18062	Human cDNA sequenc
c 525	20	1.9	657	22	AAV07935	Human TDC protein-	c 598	20	1.9	1873	21	AAA39056	Human secreted pro
c 526	20	1.9	674	21	AAH05943	Human secreted pro	c 599	20	1.9	1897	21	AAA39083	Human secreted pro
527	20	1.9	676	21	AAH02487	Human pancreatic d	600	20	1.9	1916	22	AAH44418	Human dihydroorota
528	20	1.9	691	22	AAH03326	Human cDNA clone (	c 601	20	1.9	1949	22	AAH17260	Human cDNA sequenc
c 529	20	1.9	700	22	AAH92015	Human inflammatory	c 602	20	1.9	1954	22	AAH17532	Human cDNA sequenc
c 530	20	1.9	700	22	AAH92368	Human inflammatory	603	20	1.9	1984	22	AAH16531	Human cDNA sequenc
c 531	20	1.9	700	22	AAH92369	Human inflammatory	c 604	20	1.9	1984	22	AAH13890	Sequence flanking
532	20	1.9	700	22	AAH92682	Human inflammatory	c 605	20	1.9	2040	17	AAH3949	Fucosyltransferase
533	20	1.9	705	22	AAH07356	Human cDNA clone (	c 606	20	1.9	2042	18	AAH80111	Glycosyltransferas
534	20	1.9	719	20	AAH03213	Human cDNA clone (	c 607	20	1.9	2043	15	AAQ13330	DNA encoding a gly
535	20	1.9	740	20	AAH08412	EST clone EK491.	c 608	20	1.9	2043	18	AAH76769	Human alpha(1,3/4
c 536	20	1.9	748	20	AAZ15062	Human gene express	c 609	20	1.9	2043	18	AAH161675	Human secreted pro
c 537	20	1.9	748	22	AAZ24829	Human cDNA clone (	c 610	20	1.9	2043	18	AAH161675	Human polynucleoti
c 538	20	1.9	748	22	AAH08166	Human secreted pro	c 611	20	1.9	2065	22	AAH05414	Human cDNA sequenc
c 539	20	1.9	779	21	AAH059174	Human cDNA clone (	c 612	20	1.9	2070	22	AAH14779	Human cDNA sequenc
540	20	1.9	796	20	AAH37484	Human cDNA clone (	613	20	1.9	2116	22	AAH13650	Human cDNA sequenc
c 541	20	1.9	799	22	AAH04675	Human cDNA clone (	614	20	1.9	2135	22	AAH16576	Human cDNA sequenc
c 542	20	1.9	804	22	AAH06560	Human cDNA clone (	615	20	1.9	2148	22	AAH18693	Human cDNA sequenc
543	20	1.9	811	22	AAH73279	Human cervical can	c 616	20	1.9	2152	22	AAH18693	Polynucleotide seq
c 544	20	1.9	847	22	AAH08516	Human secreted pro	617	20	1.9	2156	20	AAH18087	Human cDNA sequenc
c 545	20	1.9	847	22	AAH08516	Human secreted pro	c 618	20	1.9	2156	22	AAH18087	DNA ligase 10 cDNA
c 546	20	1.9	919	20	AAH05451	Human secreted pro	c 619	20	1.9	2209	22	AAH18440	Human cDNA sequenc
c 547	20	1.9	942	20	AAH61340	DNA encoding a hum	c 620	20	1.9	2246	22	AAH16261	Human cDNA sequenc
c 548	20	1.9	942	20	AAH81838	Human malignancy-a	c 621	20	1.9	2266	22	AAH18640	Human cDNA sequenc
c 549	20	1.9	960	21	AAH79789	Human secreted pro	c 622	20	1.9	2296	22	AAH60086	Human polynucleoti
c 550	20	1.9	968	22	AAH98581	Human EST-derived	c 623	20	1.9	2316	22	AAH05468	Human secreted pro
c 551	20	1.9	976	22	AAH63545	Human kidney relat	c 624	20	1.9	2326	22	AAH85833	DNA encoding human
c 552	20	1.9	992	21	AAH59252	Human secreted pro	c 625	20	1.9	2326	22	AAH31279	Human mammatatin
c 553	20	1.9	1001	21	AAH57825	Human secreted pro	c 626	20	1.9	2351	22	AAH27259	Human secreted pro
c 554	20	1.9	1009	22	AAH57825	Human colorectal c	c 627	20	1.9	2351	22	AAH17651	Human cDNA sequenc
c 555	20	1.9	1017	20	AAH97756	Human breast and o	c 628	20	1.9	2355	22	AAH31280	Human mammatatin
c 556	20	1.9	1028	21	AAH21761	Human secreted pro	c 629	20	1.9	2380	19	AAV34456	Human MHC class I
557	20	1.9	1086	21	AAH74264	Human secreted pro	c 630	20	1.9	2400	22	AAH15537	Human cDNA sequenc
c 558	20	1.9	1138	21	AAH79743	Human secreted pro	c 631	20	1.9	2417	19	AAV24305	Human mammary cell
c 559	20	1.9	1160	20	AAH84540	Human secreted pro	c 632	20	1.9	2418	22	AAH17392	Human cDNA sequenc
c 560	20	1.9	1162	21	AAH59157	Human secreted pro	c 633	20	1.9	2418	22	AAH31278	Human cDNA sequenc
c 561	20	1.9	1222	22	AAH94385	Human foetal cDNA,	c 634	20	1.9	2498	22	AAH18310	Human mammatatin
c 562	20	1.9	1227	22	AAH45078	Human secreted pro	c 635	20	1.9	2739	22	AAH17762	Human cDNA sequenc
563	20	1.9	1251	21	AAH59177	Human secreted pro	c 636	20	1.9	2777	19	AAH69329	Human LIR-18A3 cDN
c 564	20	1.9	1252	21	AAH22419	Human secreted pro	c 637	20	1.9	2777	21	AAH54592	Leukocyte immunogl
c 565	20	1.9	1320	21	AAH22398	Human secreted pro	c 638	20	1.9	2790	19	AAH38608	Human monocyte gen
c 566	20	1.9	1321	22	AAH19195	Human secreted pro	c 639	20	1.9	2823	22	AAH16591	Human cDNA sequenc
c 567	20	1.9	1331	22	AAH59595	Human secreted pro	c 640	20	1.9	2913	22	AAH14622	Human LIR-P3G2 cDN
c 568	20	1.9	1334	22	AAH91877	Human cell cycle a	c 641	20	1.9	2922	19	AAH69328	Human cDNA sequenc
c 569	20	1.9	1380	22	AAH44674	Novel protein kina	c 642	20	1.9	2922	21	AAH54591	Leukocyte immunogl
c 570	20	1.9	1413	21	AAH00684	Human secreted pro	643	20	1.9	2950	15	AAQ58744	aMEF2. Homo sapie
571	20	1.9	1419	22	AAH60860	Human polynucleoti	644	20	1.9	2958	15	AAQ58740	MEF2. Homo sapie
c 572	20	1.9	1425	22	AAH64215	Human secreted pro	c 645	20	1.9	3051	22	AAH17317	Human cDNA sequenc
c 573	20	1.9	1426	22	AAH159074	Human secreted pro	646	20	1.9	3137	22	AAH17317	Human cDNA sequenc
c 574	20	1.9	1470	14	AAQ47355	Human polynucleoti	c 647	20	1.9	3170	22	AAH17301	Human cDNA sequenc
c 575	20	1.9	1485	22	AAH17214	Human cDNA sequenc	c 648	20	1.9	3198	22	AAH17881	Human cDNA sequenc
576	20	1.9	1559	16	AAQ02714	Myotonic dystrophy	649	20	1.9	3240	22	AAH64150	Part of human lipo
577	20	1.9	1559	20	AAZ07987	MART-1 melanoma an	c 650	20	1.9	3332	21	AAH288392	Human ZAP-3 encodi
578	20	1.9	1559	20	AAZ20065	Human melanoma ant	c 651	20	1.9	3337	21	AAH39857	Human secreted pro
579	20	1.9	1561	22	AAH13688	Human cDNA sequenc	580	20	1.9	3337	21	AAH78421	Human secreted pro
c 580	20	1.9	1579	22	AAH17258	Human cDNA sequenc	581	20	1.9	3345	20	AAH18166	Human intestinal p
582	20	1.9	1617	22	AAH15286	Human cDNA sequenc	c 652	20	1.9	3345	22	AAH129511	Human secreted pro
583	20	1.9	1643	21	AAH87754	Human encoding huma	c 653	20	1.9	3415	21	AAH59802	Human secreted pro
c 584	20	1.9	1657	22	AAH19175	Human secreted pro	c 654	20	1.9	3415	21	AAH59802	Human secreted pro
c 585	20	1.9	1679	22	AAH17499	Human cDNA sequenc	c 655	20	1.9	3568	20	AAV99924	Human secreted pro
c 586	20	1.9	1687	15	AAQ54642	Human A2b adenosin	c 656	20	1.9	3597	21	AAH20977	Human secreted pro
c 587	20	1.9	1687	21	AAH20857	Human adenosine A2	c 657	20	1.9	3597	21	AAH34855	Human secreted pro
c 588	20	1.9	1687	21	AAH20868	Human adenosine re	c 658	20	1.9	3612	22	AAH98579	Human secreted pro
c 589	20	1.9	1687	21	AAH34735	Human adenosine re	c 659	20	1.9	3612	22	AAH57908	Human secreted pro
c 590	20	1.9	1687	21	AAH34745	Human cDNA sequenc	c 660	20	1.9	3621	22	AAH59694	Human polynucleoti
591	20	1.9	1696	22	AAH14703	Human cDNA sequenc	c 661	20	1.9	3621	19	AAH33072	Dendritic cell-der
c 592	20	1.9	1737	22	AAH94472	Human cDNA sequenc	662	20	1.9	3666	22	AAH18651	Human cDNA sequenc
c 593	20	1.9	1756	22	AAH15180	Human cDNA sequenc	c 663	20	1.9	4035	22	AAH16648	Human cDNA sequenc
c 594	20	1.9	1762	22	AAH08495	Human secreted pro	c 664	20	1.9	4459	20	AAV99094	DNA methyltransfer
c 595	20	1.9	1772	22	AAH18437	Human cDNA sequenc	c 665	20	1.9	4792	22	AAH63544	Human kidney relat
	20	1.9					c 666	20	1.9	4941	22	AAH163251	Human kidney relat
	20	1.9					c 667	20	1.9	5243	22	AAH18240	Human cDNA sequenc

c 669	20	1.9	5446	22	AAI63390	Human kidney relat
c 670	20	1.9	5830	22	AAI63546	Human kidney relat
c 671	20	1.9	6045	22	AAH18244	Human cDNA sequenc
c 672	20	1.9	6210	15	AAQ63815	c-fos gene. Homo
c 673	20	1.9	6210	21	AAF21091	Human low adenosin
c 674	20	1.9	6210	22	AAA34969	Human adenosine re
c 675	20	1.9	6235	21	AAH18663	Human cDNA sequenc
c 676	20	1.9	6235	22	AAZ29169	Human G-CSF genom
c 677	20	1.9	6312	22	AAI62822	Human genomic DNA
c 678	20	1.9	6679	21	AAZ29170	Partial sequence o
c 679	20	1.9	6711	21	AAF21122	Human low adenosin
c 680	20	1.9	6711	21	AAA35000	Human adenosine re
c 681	20	1.9	6799	22	AAI58419	Human polynucleoti
c 682	20	1.9	7144	21	AAF21439	Human adenosine re
c 683	20	1.9	7240	22	AAH50570	Insulin receptor g
c 684	20	1.9	7240	22	AAH31170	Human insulin rece
c 685	20	1.9	7240	22	AAH31267	Human insulin rece
c 686	20	1.9	7240	22	AAH31267	Human insulin rece
c 687	20	1.9	7397	19	AAH60578	Human tumour suppr
c 688	20	1.9	7737	22	AAI62831	Human genomic DNA
c 689	20	1.9	8205	22	AAI62831	Methods for diagno
c 690	20	1.9	8353	18	AAI62948	Human low adenosin
c 691	20	1.9	9203	21	AAI62948	Human adenosine re
c 692	20	1.9	9203	21	AAI62948	Human adenosine re
c 693	20	1.9	9203	21	AAI62948	Human adenosine re
c 694	20	1.9	9203	21	AAI62948	Human adenosine re
c 695	20	1.9	9203	21	AAI62948	Human adenosine re
c 696	20	1.9	9203	21	AAI62948	Human adenosine re
c 697	20	1.9	9203	21	AAI62948	Human adenosine re
c 698	20	1.9	9203	21	AAI62948	Human adenosine re
c 699	20	1.9	9203	21	AAI62948	Human adenosine re
c 700	20	1.9	9203	21	AAI62948	Human adenosine re
c 701	20	1.9	9203	21	AAI62948	Human adenosine re
c 702	20	1.9	9203	21	AAI62948	Human adenosine re
c 703	20	1.9	9203	21	AAI62948	Human adenosine re
c 704	20	1.9	9203	21	AAI62948	Human adenosine re
c 705	20	1.9	9203	21	AAI62948	Human adenosine re
c 706	20	1.9	9203	21	AAI62948	Human adenosine re
c 707	20	1.9	9203	21	AAI62948	Human adenosine re
c 708	20	1.9	9203	21	AAI62948	Human adenosine re
c 709	20	1.9	9203	21	AAI62948	Human adenosine re
c 710	20	1.9	9203	21	AAI62948	Human adenosine re
c 711	20	1.9	9203	21	AAI62948	Human adenosine re
c 712	20	1.9	9203	21	AAI62948	Human adenosine re
c 713	20	1.9	9203	21	AAI62948	Human adenosine re
c 714	20	1.9	9203	21	AAI62948	Human adenosine re
c 715	20	1.9	9203	21	AAI62948	Human adenosine re
c 716	20	1.9	9203	21	AAI62948	Human adenosine re
c 717	20	1.9	9203	21	AAI62948	Human adenosine re
c 718	20	1.9	9203	21	AAI62948	Human adenosine re
c 719	20	1.9	9203	21	AAI62948	Human adenosine re
c 720	20	1.9	9203	21	AAI62948	Human adenosine re
c 721	20	1.9	9203	21	AAI62948	Human adenosine re
c 722	20	1.9	9203	21	AAI62948	Human adenosine re
c 723	20	1.9	9203	21	AAI62948	Human adenosine re
c 724	20	1.9	9203	21	AAI62948	Human adenosine re
c 725	20	1.9	9203	21	AAI62948	Human adenosine re
c 726	20	1.9	9203	21	AAI62948	Human adenosine re
c 727	20	1.9	9203	21	AAI62948	Human adenosine re
c 728	20	1.9	9203	21	AAI62948	Human adenosine re
c 729	20	1.9	9203	21	AAI62948	Human adenosine re
c 730	20	1.9	9203	21	AAI62948	Human adenosine re
c 731	20	1.9	9203	21	AAI62948	Human adenosine re
c 732	20	1.9	9203	21	AAI62948	Human adenosine re
c 733	20	1.9	9203	21	AAI62948	Human adenosine re
c 734	20	1.9	9203	21	AAI62948	Human adenosine re
c 735	20	1.9	9203	21	AAI62948	Human adenosine re
c 736	20	1.9	9203	21	AAI62948	Human adenosine re
c 737	20	1.9	9203	21	AAI62948	Human adenosine re
c 738	20	1.9	9203	21	AAI62948	Human adenosine re
c 739	20	1.9	9203	21	AAI62948	Human adenosine re
c 740	20	1.9	9203	21	AAI62948	Human adenosine re
c 741	20	1.9	9203	21	AAI62948	Human adenosine re

c 742	20	1.9	24031	17	AAI17525	Mutated BRCA1 geno
c 743	20	1.9	24218	22	AAI62935	Human genomic DNA
c 744	20	1.9	26006	22	AAI62935	Human breast or ov
c 745	20	1.9	26006	22	AAI62935	Human kidney relat
c 746	20	1.9	26013	22	AAI62590	Human breast or ov
c 747	20	1.9	26013	22	AAI62590	Human kidney relat
c 748	20	1.9	26018	22	AAI62591	Human breast or ov
c 749	20	1.9	26018	22	AAI62591	Human kidney relat
c 750	20	1.9	26328	20	AAZ32184	Human prothrombin
c 751	20	1.9	32152	22	AAI57791	Human colorectal c
c 752	20	1.9	32183	22	AAI63617	Human kidney relat
c 753	20	1.9	32204	22	AAI57790	Human colorectal c
c 754	20	1.9	33780	22	AAH24652	Nucleotide sequenc
c 755	20	1.9	34488	22	AAI62935	Human neuroblastom
c 756	20	1.9	38186	22	AAI62935	Human neuroblastom
c 757	20	1.9	38186	22	AAI62935	Human neuroblastom
c 758	20	1.9	39198	22	AAI62935	Human neuroblastom
c 759	20	1.9	44453	20	AAI62935	Human polyamine-mo
c 760	20	1.9	49999	20	AAI62935	Human kidney amino
c 761	20	1.9	49999	20	AAI62935	Human LOBO homolog
c 762	20	1.9	49999	20	AAI62935	Human LOBO homolog
c 763	20	1.9	50000	21	AAI62935	Polymorphic repeat
c 764	20	1.9	50000	21	AAI62935	Nucleotide sequenc
c 765	20	1.9	51474	22	AAI62935	Nucleotide sequenc
c 766	20	1.9	52216	22	AAI62935	Nucleotide sequenc
c 767	20	1.9	72928	20	AAI62935	Human neuroblastom
c 768	20	1.9	72928	20	AAI62935	Human neuroblastom
c 769	20	1.9	80331	22	AAI62935	Human ASH1J 5' ge
c 770	20	1.9	80331	22	AAI62935	Human ASH1J 5' ge
c 771	20	1.9	81001	22	AAI62935	Human histone deac
c 772	20	1.9	81145	22	AAI62935	Human histone deac
c 773	20	1.9	84607	20	AAI62935	Human apolipoprote
c 774	20	1.9	90050	21	AAI62935	Human PACAP genom
c 775	20	1.9	97662	22	AAI62935	Wild type (C57BL/6
c 776	20	1.9	110000	22	AAI62935	Genomic sequence o
c 777	20	1.9	114793	22	AAI62935	Nucleotide sequenc
c 778	20	1.9	114793	22	AAI62935	Human genome from
c 779	20	1.9	117609	21	AAI62935	Human receptor-rel
c 780	20	1.9	162450	21	AAI62935	Retinoblastoma bin
c 781	20	1.9	235033	19	AAI62935	Hereditary haemoch

## ALIGNMENTS

RESULT 1

ID AAD03816 standard; cDNA; 1041 BP.

AC AAD03816;

DT 19-JUN-2001 (first entry)

DE Human Kinase cDNA #5.

KW Human: kinase; gene therapy; bioreactor; mental disorder;

XX Homo sapiens.

XX Key

XX CDS

XX Location/Qualifiers

XX 1..1041

XX /\*tag= a

XX /product= "Human kinase #5"

XX /note= "The coding region does not include stop codon"

XX /partial

XX WO200123579-A1.

XX 05-APR-2001.

XX 27-SEP-2000; 2000WO-US26621.

XX

XX

XX

XX

PR 28-SEP-1999; 99US-0156511.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI: 2001-266166/27.  
DR P-PSDB; AAE00494.  
XX  
XX New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases  
XX  
XX Claim 1; Page 31-32; 38pp; English.  
XX  
XX The present sequence is a cDNA encoding novel human protein  
CC (NHP) known as human kinase. The human kinases share structural  
CC similarity with animal kinases. More particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.  
XX  
XX Sequence 1041 BP; 344 A; 190 C; 228 G; 279 T; 0 other;  
SQ

Query Match 100.0%; Score 1041; DB 22; Length 1041;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggaagaatgaataatagctaaactggaagaaggtcttatgggtgtattcaaa 60  
Db 1 atggaagaatgaataatagctaaactggaagaaggtcttatgggtgtattcaaa 60  
Qy 61 tgcagaaacaaacctctggacaagtagtagctgttaaaaaattgtggaaatcgaagat 120  
Db 61 tgcagaaacaaacctctggacaagtagtagctgttaaaaaattgtggaaatcgaagat 120  
Qy 121 gatctgtgttaaaaaatagcactaagaagaatacgtatgttgaagcaattaaacat 180  
Db 121 gatctgtgttgaagaataatagcactaagaagaatacgtatgttgaagcaattaaacat 180  
Qy 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaataatcattagtttt 240  
Db 181 ccaaatcttgtgaacctcatcgagggtgttcaggagaaaaaggaataatcattagtttt 240  
Qy 241 gaatactgtgatacacaacttttaaatgagctggaaagaaacccaatggaggtgtgtat 300  
Db 241 gaatactgtgatacacaacttttaaatgagctggaaagaaacccaatggaggtgtgtat 300  
Qy 301 ggagtgatacaaaagcgtattatggcaaacacttcaagcttcttaattctgtcatatacat 360  
Db 301 ggagtgatacaaaagcgtattatggcaaacacttcaagcttcttaattctgtcatatacat 360  
Qy 361 aactgtattcacagagataaaaaacctgaaatatcttaataactaagcaaggaataatc 420  
Db 361 aactgtattcacagagataaaaaacctgaaatatcttaataactaagcaaggaataatc 420

Qy 421 aagatttgacttcgggtttgcacaaattctgattccaggagatgctcacaccgattat 480  
Db 421 aagatttgacttcgggtttgcacaaattctgattccaggagatgctcacaccgattat 480  
Qy 481 gtactgaagaatgtaccagactcctgaactcttctgtgggagatactcagtagttct 540  
Db 481 gtactgaagaatgtaccagactcctgaactcttctgtgggagatactcagtagttct 540  
Qy 541 tcagtcgatatatgggctattgtgttttttcagagactcctcagagccagccactg 600  
Db 541 tcagtcgatatatgggctattgtgttttttcagagactcctcagagccagccactg 600  
Qy 601 tgccttgaaaaatcagatgtggaccaactttatctgataatcagacaactagtagagacg 660  
Db 601 tgccttgaaaaatcagatgtggaccaactttatctgataatcagacaactagtagagacg 660  
Qy 661 gggtttcgcatgttgaccagggtgtctcgaactcttgacgtcaagtgtaccctgcc 720  
Db 661 gggtttcgcatgttgaccagggtgtctcgaactcttgacgtcaagtgtaccctgcc 720  
Qy 721 gtacgctctcaaaagtgtcgaattacaggaataattatcccaagacatcaatcaatttt 780  
Db 721 gtacgctctcaaaagtgtcgaattacaggaataattatcccaagacatcaatcaatttt 780  
Qy 781 aaaaagtaacgggtttttccatgcacatcagatcactcagcagcaagacacatggaactctt 840  
Db 781 aaaaagtaacgggtttttccatgcacatcagatcactcagcagcaagacacatggaactctt 840  
Qy 841 gaggaagaagtctcagatgttcatctctgtggtctgaaactcgaaggggtgtctgaag 900  
Db 841 gaggaagaagtctcagatgttcatctctgtggtctgaaactcgaaggggtgtctgaag 900  
Qy 901 atgaatccagatgacagattaaacctgtcccaactcctggagagctcactcttgattct 960  
Db 901 atgaatccagatgacagattaaacctgtcccaactcctggagagctcactcttgattct 960  
Qy 961 ttccaagaggcccaaatataaagaaagcagctaatgaagaaagaaagaaagacgacaa 1020  
Db 961 ttccaagaggcccaaatataaagaaagcagctaatgaagaaagaaagaaagacgacaa 1020  
Qy 1021 caggtactctcgtctcaaaagt 1041  
Db 1021 caggtactctcgtctcaaaagt 1041

RESULT 2  
AAD03813  
ID AAD03813 standard; cDNA: 1068 BP.  
XX  
AC AAD03813;  
XX  
DT 19-JUN-2001 (first entry)  
XX  
DE Human kinase cDNA #2.  
XX  
KW Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 1..1068  
FT /\*tag= a  
FT /product= "Human kinase #2"  
FT /note= "The coding region does not include stop codon"  
FT /partial  
XX  
XX WO200123579-A1.  
XX PN  
XX 05-APR-2001.  
XX PD  
XX 27-SEP-2000; 2000WO-US26621.  
XX PF







CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.  
XX  
SQ Sequence 972 BP; 323 A; 172 C; 214 G; 263 T; 0 other;

Query Match 62.6%; Score 652; DB 22; Length 972;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaagaagtataaaattagctaaagactggagaagggtcttatgggttgattcaaa 60  
Db |||||||  
QY 1 atggaagaagtataaaattagctaaagactggagaagggtcttatgggttgattcaaa 60  
Db |||||||  
QY 61 tgcagaaacaaacctctgcacagtaagctagctgttaaaaaatttgggaatctgaagat 120  
Db |||||||  
QY 61 tgcagaaacaaacctctgcacagtaagctagctgttaaaaaatttgggaatctgaagat 120  
Db |||||||  
QY 121 gatctgtgttaagaaatagcactaagagaatacgtatgttggaagcaattaaacat 180  
Db |||||||  
QY 121 gatctgtgttaagaaatagcactaagagaatacgtatgttggaagcaattaaacat 180  
Db |||||||  
QY 181 ccaaatctgtgaacctcatcaggtgttcagagaaaaagaaatgaattagttttt 240  
Db |||||||  
QY 181 ccaaatctgtgaacctcatcaggtgttcagagaaaaagaaatgaattagttttt 240  
Db |||||||  
QY 241 gaatactgtatcatcaccttttaaatgagctggaaagaaacccaaatggagtgtctgat 300  
Db |||||||  
QY 241 gaatactgtatcatcaccttttaaatgagctggaaagaaacccaaatggagtgtctgat 300  
Db |||||||  
QY 301 ggaatgtcaaaagcgtattatggcaaacacttcaagctcttaattcttgcataacat 360  
Db |||||||  
QY 301 ggaatgtcaaaagcgtattatggcaaacacttcaagctcttaattcttgcataacat 360  
Db |||||||  
QY 361 aactgtattcacagagatataaaacctgaaatattcttaataactaagcaaggaataatc 420  
Db |||||||  
QY 361 aactgtattcacagagatataaaacctgaaatattcttaataactaagcaaggaataatc 420  
Db |||||||  
QY 421 aagatttgacttcgggtttgcacaaattctgattccagagatgcctacacagattat 480  
Db |||||||  
QY 421 aagatttgacttcgggtttgcacaaattctgattccagagatgcctacacagattat 480  
Db |||||||  
QY 481 gtactacagatgtaccagctcctgaactctctgttggaagatactcagtatggttct 540  
Db |||||||  
QY 481 gtactacagatgtaccagctcctgaactctctgttggaagatactcagtatggttct 540  
Db |||||||  
QY 541 tcagtcgatatgggtattgtgtgttttgcagagctcctgacagggccagccactg 600  
Db |||||||  
QY 541 tcagtcgatatgggtattgtgtgttttgcagagctcctgacagggccagccactg 600  
Db |||||||  
QY 601 tggcctggaaaatcagatgtggaccacatttctatgataatcagaaactag 652  
Db |||||||  
QY 601 tggcctggaaaatcagatgtggaccacatttctgataatcagaaactag 652  
Db |||||||

RESULT 5  
AAD03818  
ID AAD03818 standard; cDNA; 1819 BP.  
XX  
AC AAD03818;  
XX  
DT 19-JUN-2001 (first entry)

XX Human kinase cDNA #7.  
DE  
XX  
KW Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder; polymorphism; ss.  
XX Homo sapiens.  
OS  
PH Key Location/Qualifiers  
FT 5'UTR 1..137  
FT /\*tag= a  
FT CDS 138..701  
FT /\*tag= b  
FT 3'UTR /product= "Human kinase #1"  
FT 702..1819  
FT /\*tag= c  
FT misc\_difference 1684  
FT /\*tag= d  
FT /note= "polymorphism site; given as r in the sequence"  
XX WO200123579-A1.  
PN  
XX  
XX  
XX  
XX 05-APR-2001.  
XX  
XX 27-SEP-2000; 2000WO-US26621.  
XX  
XX 28-SEP-1999; 99US-0156511.  
XX  
XX (LEXI-) LEXICON GENETICS INC.  
XX  
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
XX  
XX WPI; 2001-266166/27.  
XX  
XX P-PSDB; AAE00490.  
XX  
XX New isolated human kinase polynucleotide useful for generating  
XX antibodies, as reagents in diagnostic assays and for screening for  
XX compounds useful for treating mental, biological or medical diseases -  
XX  
XX Disclosure; Page 34-35; 38pp; English.  
XX  
XX The present sequence encodes a novel human protein (NHP) which  
XX is a human kinase. A polymorphism was identified in the 3' UTR  
XX of the present sequence. The human kinases share structural  
XX similarity with animal kinases, more particularly serine or  
XX threonine protein kinases. Human kinase cDNA is useful for the  
XX detection of mutant human kinase for the diagnosis of disease,  
XX and also as a therapeutic. It is useful for screening drugs  
XX effective in the treatment of symptomatic or phenotypic  
XX manifestations perturbing the normal function of NHP in the  
XX body. The NHP nucleotide sequences are useful for generation of  
XX antibodies, as reagents in diagnostic assays, for the  
XX identification of other cellular gene products related to human  
XX kinases, and as reagents in assays for screening compounds that  
XX are useful for treating mental, biological or medical disorders.  
XX NHP oligonucleotides are used as probes. The labelled NHP probes  
XX are useful for screening human genomic library for identifying  
XX mutations within the exons, introns and splice sites that can  
XX be used in diagnostics and pharmacogenomics. Nucleotide construct  
XX encoding NHP products are used to genetically engineer cells  
XX in vivo that functions as bioreactors in the body delivering a  
XX continuous supply of NHP to the body. Nucleotide constructs  
XX encoding functional NHPs are used in gene therapy for the  
XX modulation of NHP expression.  
SQ Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;

Query Match 54.9%; Score 572; DB 22; Length 1819;  
Best Local Similarity 100.0%; Pred. No. 2,9e-276;  
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 attccagagagatgctacaccgattatgactacagagatgctaccgagctcctgaactt 513  
DB 897 attccagagagatgctacaccgattatgactacagagatgctaccgagctcctgaactt 956  
QY 514 ctgtgtggagagatcactcagtgatgttcttcagtcgatatatgggtcatttgggtgttttt 573  
DB 957 ctgtgtggagagatcactcagtgatgttcttcagtcgatatatgggtcatttgggtgttttt 1016  
QY 574 gcagagctcctcagacgaccacactgtgctgtggaatacagatgtggacacactttat 633  
DB 1017 gcagagctcctcagacgaccacactgtgctgtggaatacagatgtggacacactttat 1076  
QY 634 ctgataatcagacactagtagagagcgggtttccgcatgtttaccagcgttgctcgaa 693  
DB 1077 ctgataatcagacactagtagagagcgggtttccgcatgtttaccagcgttgctcgaa 1136  
QY 694 cctctgacgtcgaagtacacactcgcgtagcctctcaagtcgtggaattacagggaaa 753  
DB 1137 cctctgacgtcgaagtacacactcgcgtagcctctcaagtcgtggaattacagggaaa 1196  
QY 754 ttaatccagacatcaatcaatctttaaaagtaacgggtttttccatggcatacagata 813  
DB 1197 ttaatccagacatcaatcaatctttaaaagtaacgggtttttccatggcatacagata 1256  
QY 814 cctgagccagagacatggaactcttgagaaaaagttctcagatgttcaatcctgtggt 873  
DB 1257 cctgagccagagacatggaactcttgagaaaaagttctcagatgttcaatcctgtggt 1316  
QY 874 ctgaactcctgaaggggtctgaagataatccagatgacagattaaacctgttcccaa 933  
DB 1317 ctgaactcctgaaggggtctgaagataatccagatgacagattaaacctgttcccaa 1376  
QY 934 ctcctggagagctcctactttgattcttttcaagaggcccaattaaagaaagcaggt 993  
DB 1377 ctcctggagagctcctactttgattcttttcaagaggcccaattaaagaaagcaggt 1436  
QY 994 aatgaagaagaacagaagacgccaacaggt 1025  
DB 1437 aatgaagaagaacagaagacgccaacaggt 1468  
RESULT 6  
AAD03812  
ID AAD03812 standard; cDNA; 561 BP.  
XX AC AAD03812;  
XX DT 19-JUN-2001 (first entry)  
XX DE Human kinase cDNA #1.  
XX KW Human; kinase; gene therapy; bioreactor; mental disorder;  
XX KW biological disorder; ss.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX FT CDS 1..561  
XX FT /\*tag= a  
XX FT /\*product= "Human Kinase #1"  
XX FT /\*note= "The coding region does not include stop codon"  
XX FT /partial  
XX PN WO200123579-A1.  
XX PD 05-APR-2001.  
XX PF 27-SEP-2000; 2000WO-US26621.  
XX PR 28-SEP-1999; 99US-0156511.  
XX PA (LEXI-) LEXICON GENETICS INC.

PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI: 2001-266166/27.  
DR P-PSDB; AAE00490.  
XX New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases -  
XX Disclosure; Page 27; 38pp; English.  
XX The present sequence is a cDNA encoding novel human protein  
CC (NHP) known as human kinase. The human kinases share structural  
CC similarity with animal kinases, more particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.  
XX Sequence 561 BP; 189 A; 87 C; 112 G; 173 T; 0 other;

Query Match 43.6%; Score 454; DB 22; Length 561;  
Best Local Similarity 100.0%; Pred. No. 3.1e-217;  
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atggaaaagtataaaaattagctaagactggagaggggtctatgggtgtattcaaa 60  
DB 1 atggaaaagtataaaaattagctaagactggagaggggtctatgggtgtattcaaa 60  
QY 61 tgcagaaaacaaacctctgacaaagtagtagctgttaaataattgtggaatctgaagat 120  
DB 61 tgcagaaaacaaacctctgacaaagtagtagctgttaaataattgtggaatctgaagat 120  
QY 121 gatcctgtgttaagaaaaatagcaactagagaaaatagctatgttgaagcaattaaaacat 180  
DB 121 gatcctgtgttaagaaaaatagcaactagagaaaatagctatgttgaagcaattaaaacat 180  
QY 181 ccaaatctgtgaacctcactcaggtgttccaggagaaaaagaaaatgcattagtgtttt 240  
DB 181 ccaaatctgtgaacctcactcaggtgttccaggagaaaaagaaaatgcattagtgtttt 240  
QY 241 gaatactgtgatcatacaacttttaaatgagctggaaagaaacccaatggagttgctgat 300  
DB 241 gaatactgtgatcatacaacttttaaatgagctggaaagaaacccaatggagttgctgat 300  
QY 301 ggaagtatcaaaaagcgtattattatggcaaacacttcaagctcttaattctgtcatacat 360  
DB 301 ggaagtatcaaaaagcgtattattatggcaaacacttcaagctcttaattctgtcatacat 360  
QY 361 aactgtattcacagagataataaacctgaaatattcttaataactaagcaaggaataatc 420  
DB 361 aactgtattcacagagataataaacctgaaatattcttaataactaagcaaggaataatc 420  
QY 421 aagatttgtgacttcgggttttgcaaaaattctga 454  
DB 421 aagatttgtgacttcgggttttgcaaaaattctga 454

RESULT 7  
 AAD03815  
 ID AAD03815 standard; cDNA; 594 BP.  
 XX  
 AC AAD03815;  
 XX  
 DT 19-JUN-2001 (first entry)  
 XX  
 DE Human kinase cDNA #4.  
 XX  
 KW Human; kinase; gene therapy; bioreactor; mental disorder;  
 KW biological disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..594  
 FT FT /\*tag= a  
 FT FT /product= "Human kinase #4"  
 FT FT /note= "The coding region does not include stop codon"  
 FT FT /partial  
 XX  
 PN WO200123579-A1.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 27-SEP-2000; 2000WO-US26621.  
 XX  
 PR 28-SEP-1999; 99US-0156511.  
 XX  
 PA (LEXI-) LEXICON GENETICS INC.  
 XX  
 PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
 XX  
 DR WPI; 2001-266166/27.  
 DR P-PSDB; AAE00493.  
 XX  
 PT New isolated human kinase polynucleotide useful for generating  
 PT antibodies, as reagents in diagnostic assays and for screening for  
 PT compounds useful for treating mental, biological or medical diseases -  
 PS Disclosure; Page 31; 38pp; English.  
 XX  
 CC The present sequence is a cDNA encoding novel human protein  
 CC (NHP) known as human kinase. The human kinases share structural  
 CC similarity with animal kinases. The human kinases share structural  
 CC threonine protein kinases. more particularly serine or  
 CC detection of mutant human kinase for the diagnosis of disease,  
 CC and also as a therapeutic. It is useful for screening drugs  
 CC effective in the treatment of symptomatic or phenotypic  
 CC manifestations perturbing the normal function of NHP in the  
 CC body. The NHP nucleotide sequences are useful for generation of  
 CC antibodies, as reagents in diagnostic assays, for the  
 CC identification of other cellular gene products related to human  
 CC kinases, and as reagents in assays for screening compounds that  
 CC are useful for treating mental, biological or medical disorders.  
 CC NHP oligonucleotides are used as probes. The labelled NHP probes  
 CC are useful for screening human genomic library for identifying  
 CC polymorphisms and as primers in amplification assays to detect  
 CC mutations within the exons, introns and splice sites that can  
 CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
 CC encoding NHP products are used to genetically engineer cells  
 CC in vivo that functions as bioreactors in the body delivering a  
 CC continuous supply of NHP to the body. Nucleotide constructs  
 CC encoding functional NHPs are used in gene therapy for the  
 CC modulation of NHP expression.  
 XX  
 SQ Sequence 594 BP; 203 A; 94 C; 115 G; 182 T; 0 other;

Query Match

43.6%; Score 454; DB 22; Length 594;

Best Local Similarity 100.0%; Pred. No. 3.1e-217;  
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 atggaagatgataaaatagctaaagctggaagggctcttatgggtgtatcaca 60  
 Db 1 atggaagatgataaaatagctaaagctggaagggctcttatgggtgtatcaca 60  
 QY 61 tgcagaaacaaacctctggacaaagtagctgttaaaaaatttgggaatctgaagat 120  
 Db 61 tgcagaaacaaacctctggacaaagtagctgttaaaaaatttgggaatctgaagat 120  
 QY 121 gatctgtgttaagaaaaatagcactaagaagaaatcgtatgttgaagcaattaaaaaat 180  
 Db 121 gatctgtgttaagaaaaatagcactaagaagaaatcgtatgttgaagcaattaaaaaat 180  
 QY 181 ccaaatcttgtgaacctcatcgaggtgttcagagagaaaaaagaaatgcattagtttt 240  
 Db 181 ccaaatcttgtgaacctcatcgaggtgttcagagagaaaaaagaaatgcattagtttt 240  
 QY 241 gaatactgtgatacatcacttttaaatgagctggaagaaaccccaaatgagttgctgat 300  
 Db 241 gaatactgtgatacatcacttttaaatgagctggaagaaaccccaaatgagttgctgat 300  
 QY 301 ggagtgatcaaaagcgtattatgcaaacacttcaagctccttaatttctgcatacat 360  
 Db 301 ggagtgatcaaaagcgtattatgcaaacacttcaagctccttaatttctgcatacat 360  
 QY 361 aactgtattcacagagatataaaacctgaaaatattcttaataagcaaggaataatc 420  
 Db 361 aactgtattcacagagatataaaacctgaaaatattcttaataagcaaggaataatc 420  
 QY 421 aagattgtgacttcgggtttgcacaaattctga 454  
 Db 421 aagattgtgacttcgggtttgcacaaattctga 454

RESULT 8

AAS06725  
 ID AAS06725 standard; cDNA; 1083 BP.  
 AC AAS06725;  
 XX  
 DT 12-SEP-2001 (first entry)  
 XX  
 DE Polynucleotide sequence encoding human protein kinase #25.  
 XX  
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
 KW metabolic disorder; immune related disease; neurological disorder;  
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
 KW reproductive disorder; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200138503-A2.  
 XX  
 PD 31-MAY-2001.  
 XX  
 PF 22-NOV-2000; 2000WO-US32085.  
 XX  
 PR 24-NOV-1999; 99US-0167482.  
 XX  
 PA (SUG-) SUGEN INC.  
 XX  
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
 PI Flanagan P, Clary D;  
 XX  
 DR WPI; 2001-343950/36.  
 DR P-PSDB; AAU03525.  
 XX  
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing  
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
 PT neuronal-associated diseases, and microbial infections -  
 XX

immune disorder; cardiovascular disease; neurodegenerative disease; cancer; autoimmune disorder; stroke; inflammatory bowel disease; inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

Example 1; Figure 1; 433pp; English.  
AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel protein kinases have been identified as members of the tyrosine or serine/threonine kinase (PK and STK) families. The polynucleotides encoding protein kinases and the polypeptides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate kinase expression. For example, they may be used to treat cancers (especially cancers of haematopoietic origin), cardiovascular disease (especially atherosclerosis), metabolic disorders (e.g. diabetes), immune related diseases (e.g. rheumatoid arthritis), neurological disorders (e.g. schizophrenia), neurodegenerative disorders (e.g. Parkinson's disease), inflammatory disorders (e.g. asthma), infectious diseases (e.g. HIV) and reproductive disorders (e.g. infertility). Additionally, polynucleotides encoding protein kinases may be used for gene therapy and as DNA probes in diagnostic assays. The protein kinase polypeptides may be used as antigens in the production of antibodies against the protein kinases and in assays to identify modulators of protein kinase expression and activity.  
Sequence 1083 BP; 366 A; 199 C; 225 G; 293 T; 0 other;

Query Match 43.6%; Score 454; DB 22; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 3.2e-217;  
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atggaagaatgataaaattagctaaactgagaggaagggtcttattggtgtattcaaa 60  
DB 1 atggaagaatgataaaattagctaaactgagaggaagggtcttattggtgtattcaaa 60  
QY 61 tgcagaacaaaacctctgacaaagttagctgtgttaaaaaatttggaaactgaagat 120  
DB 61 tgcagaacaaaacctctgacaaagttagctgtgttaaaaaatttggaaactgaagat 120  
QY 121 gattctgtgttaagaaatagcactaagagaatacgtattgtgaagcaattaaacat 180  
DB 121 gattctgtgttaagaaatagcactaagagaatacgtattgtgaagcaattaaacat 180  
QY 181 ccaaatctgtgaacctctcaggtgttcagagaaaaagaaaatgcatttttt 240  
DB 181 ccaaatctgtgaacctctcaggtgttcagagaaaaagaaaatgcatttttt 240  
QY 241 gaactgtgtatcacacttttaattgaagctggaagaaacccaatgaggtgtgat 300  
DB 241 gaactgtgtatcacacttttaattgaagctggaagaaacccaatgaggtgtgat 300  
QY 301 ggaagtatacaaaagcgtattatgcaaacacttcaagctcttaatttctcatatacat 360  
DB 301 ggaagtatacaaaagcgtattatgcaaacacttcaagctcttaatttctcatatacat 360  
QY 361 aactgtattcacagagatataaaacctgaaatattctataactaagaaggaataatc 420  
DB 361 aactgtattcacagagatataaaacctgaaatattctataactaagaaggaataatc 420  
QY 421 aagatttgtgacttcgggttgcacaaattctga 454  
DB 421 aagatttgtgacttcgggttgcacaaattctga 454

RESULT 9  
ID AAF44669  
ID AAF44669 standard; cDNA; 911 BP.  
XX AAF44669;  
AC AAF44669;  
DT 27-MAR-2001 (first entry)  
XX Novel protein kinase cDNA, SEQ ID NO: 49.  
DE Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;  
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;

XX KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
XX inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.  
XX OS Homo sapiens.  
XX PN WO200073469-A2.  
XX PD 07-DEC-2000.  
XX XX 26-MAY-2000; 2000WO-US14842.  
XX PF 28-MAY-1999; 99US-0136503.  
XX XX (SUGE-) SUGEN INC.  
XX PA Plowman GD, Martinez R, Whyte D, Sudersanam S;  
PI WPI: 2001-032161/04.  
XX DR P-PSDB; AAB65642.  
XX XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers -  
XX Example 4; Fig 2; 310pp; English.  
XX The present sequence encodes a novel protein kinase. The nucleic acids  
CC and the protein kinases they encode may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression,  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC expression and activity. Diseases related to kinase expression and  
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
CC reproductive disorders.  
XX XX Sequence 911 BP; 293 A; 182 C; 178 G; 258 T; 0 other;

Query Match 35.4%; Score 368; DB 22; Length 911;  
Best Local Similarity 100.0%; Pred. No. 3.4e-174;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 aaatggagtgtctgatggatgatacaaaagcgtattatggcaaacacttcaagctcttaa 344  
DB 1 aaatggagtgtctgatggatgatacaaaagcgtattatggcaaacacttcaagctcttaa 60  
QY 345 ttctctcatatcataactgtattccacagagataaaaacctgaaaaattcttaataac 404  
DB 61 ttctctcatatcataactgtattccacagagataaaaacctgaaaaattcttaataac 120  
QY 405 taagcaagaataatcaatcaagatttgtacttcgggttgcacaaatctgtattccaggaga 464  
DB 121 taagcaagaataatcaatcaagatttgtacttcgggttgcacaaatctgtattccaggaga 180  
QY 465 tgctcacccgattatgtagctacagagatggtaaccagagctcctgaactcttctgtgggaga 524  
DB 181 tgctcacccgattatgtagctacagagatggtaaccagagctcctgaactcttctgtgggaga 240  
QY 525 tactcagtagtgttcttcagtcgatatatggctattgtgtgttttttgcagagctcct 584  
DB 241 tactcagtagtgttcttcagtcgatatatggctattgtgtgttttttgcagagctcct 300  
QY 585 gacaggccaccactgtgcttggaataatcaatgtaggaccaacttttatgtataatcag 644

```
Db 301 gacagccagcactgtgctggaatacagatgtgacacatttatctgataatcag 360
QY 645 aacactag 652
Db 361 aacactag 368

RESULT 10
AAF97873/C
ID AAF97873 standard; DNA; 12729 BP.
XX AC AAF97873;
XX AC AAF97873;
XX DT 31-MAY-2001 (first entry)
XX DE Human neuroblastoma cell line NB-1 lp36 nucleotide sequence SEQ ID NO:87.
XX DE Human; chromosome 1; lp36; neuroblastoma cell line; NB-1; anticancer;
XX KW tumour suppressor; human lp36 homozygosity deletion domain; tumour;
XX KW diagnosis; ds.
XX XX Homo sapiens.
XX OS WO200116311-A1.
XX PN 08-MAR-2001.
XX PD
XX PF 31-AUG-2000; 2000WO-JP05930.
XX PR 31-AUG-1999; 99JP-0245962.
XX PR 09-MAY-2000; 2000JP-0136266.
XX XX
XX PA (HISM ) HISAMITSU PHARM CO LTD.
XX PA (CHIB-) CHIBA PREFECTURE.
XX PI Nakagawara A;
XX XX WPI; 2001-226686/23.
XX DR
XX PT Human lp36 homozygosity deletion domain from the 36-position of first
XX PT chromosome short arm in human neuroblastoma cell lines, applicable e.g.
XX PT in gene diagnosis of tumors as well as in developing anti-cancer drugs
XX XX
XX PS Example 8; Page 213-218; 226pp; Japanese.
XX CC The present invention describes a homozygosity deletion domain
XX CC co-existing in the 36-position of the first chromosome short arm (lp36)
XX CC in human neuroblastoma. Also described are base sequences from the lp36
XX CC position of human neuroblastoma cell lines (NB-1 and MASS-NB-SCH-1), are
XX CC which are tumour suppressor genes in human neuroblastoma. The genes are
XX CC tumour suppressor genes, base sequence data of which are applicable as
XX CC tumour markers and reagents in studying mechanism of tumour body
XX CC formation, and gene diagnosis of tumours as well as in developing
XX CC anti-cancer drugs. AAF97787 to AAF97829 represent PCR primers used in
XX CC the exemplification of the present invention, and AAF97830 to AAF97874
XX CC represent sequences given in the exemplification of the present
XX XX
XX SQ Sequence 12729 BP; 4419 A; 2482 C; 2733 G; 3095 T; 0 other;
```

```
Query Match 3.8%; Score 40; DB 22; Length 12729;
Best Local Similarity 100.0%; Pred. No. 4.7e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 650 tagtagagacgggttcgccatgttgaccagcgtggtct 689
Db 1800 TAGTAGAGACGGGTTTCGCCATGTTGACCAGCGCTGCTT 1761
```

RESULT 11

```
AAC29051
ID AAC29051 standard; cDNA; 105 BP.
XX AC AAC29051;
XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 33126.
XX DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN EP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-0200610.
XX PR 26-FEB-1999; 99US-0122487.
XX PA (GEST ) GENSET.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX PS Claim 1; SEQ ID 33126; 71pp + CD-ROM; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX CC identified within the present sequence. The 5' ESTs were prepared from
XX CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX CC sequences usually correspond mainly to the 3' untranslated region (UTR)
XX CC of the mRNA because they are often obtained from oligo-dT primed cDNA
XX CC libraries. Such ESTs are not well suited for isolating cDNA sequences
XX CC derived from the 5' ends of mRNAs and even in those cases where longer
XX CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX CC They are used to obtain upstream regulatory sequences and to design
XX CC expression and secretion vectors.
XX SQ Sequence 105 BP; 17 A; 27 C; 29 G; 32 T; 0 other;
```

```
Query Match 3.2%; Score 33; DB 21; Length 105;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 669 ccatgttgaccagcgtgtgtctcgaaactcttgac 701
Db 11 ccatgttgaccagcgtgtgtctcgaaactcttgac 43
```

```
RESULT 12
AAQ61041
ID AAQ61041 standard; DNA; 411 BP.
XX AC AAQ61041;
XX XX
XX DT 16-MAR-1994 (first entry)
XX DE Human brain Expressed Sequence Tag EST01677.
XX KW Gene transcription product; genetic markers; tagging; in vivo;
XX KW transcription; mapping; locations; chromosomes; chromosomal; ss.
```

us-09-671-050-9.oli20.rng

Fri May 3 11:02:10 2002

XX OS Homo sapiens.  
XX PN W09316178-A.  
XX PD 19-AUG-1993.  
XX PF 12-FEB-1993; 93WO-US01294.  
XX PR 12-FEB-1992; 92US-0837195.  
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
XX PI Adams MD, Moreno RF, Venter CJ;  
XX WPI: 1993-272882/34.  
XX PT Enriched oligonucleotides and corresp. sequences - used as  
XX markers for human genes transcribed in-vivo, facilitate tagging  
XX of most human genes  
XX Example 4; Page 430; 500pp; English.  
XX The Expressed Sequence Tag was isolated from a human brain cDNA  
XX library as part of a large set of ESTs which can be used as markers  
XX for human genes transcribed in vivo. They can be used to facilitate  
XX tagging of most human genes, for mapping locations of expressed genes  
XX on chromosomes, for individual or forensic identification, for mapping  
XX locations of disease-associated genes, for identification of tissue  
XX type, and for prep. of antisense sequences, probes and constructs.  
XX - EST01677 has a "poor" coding probability as evaluated using the  
XX coding-region prediction program CRM. See also AAQ59041-Q61440.  
XX SQ Sequence 411 BP; 99 A; 114 C; 93 G; 105 T; 0 other;  
  
Query Match 3.2%; Score 33; DB 14; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 669 ccattgtgaccaggctggtctcgaactcttgac 701  
Db 63 ccattgtgaccaggctggtctcgaactcttgac 95  
|||||  
RESULT 13  
AAF26586/C  
ID AAF26586 standard; DNA; 2286 BP.  
XX AC AAF26586;  
XX DT 27-MAR-2001 (first entry)  
XX DE DNA encoding human secreted protein #40.  
XX KW Secreted protein; gene therapy; vaccine; cancer; leukemia;  
XX KW autoimmune disease; allergy; inflammation; graft rejection;  
XX KW hyperproliferation; cardiovascular; infection; ss.  
XX OS Homo sapiens.  
XX PN W0200076531-A1.  
XX PD 21-DEC-2000.  
XX PF 01-JUN-2000; 2000WO-US15137.  
XX PR 11-JUN-1999; 99US-0138625.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Ruben SM, Komatsoulis GA;  
XX DR

DR WPI: 2001-071148/08.  
XX Nucleic acids encoding 47 human secreted polypeptides, useful for  
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's  
XX disease and diabetic retinopathy -  
XX Claim 1; Page 467-498; 525pp; English.  
XX The present invention relates to 26 secreted human proteins. The  
XX proteins may be used in the prevention, diagnosis and treatment of  
XX diseases associated with inappropriate polypeptide expression.  
XX For example, they may be used in gene therapy or in vaccines.  
XX Typical of diseases which are potentially treatable are cancers  
XX (including leukemia), autoimmune diseases, allergies, inflammation,  
XX graft rejection, hyperproliferation, cardiovascular diseases  
XX (particularly critical limb ischemia and coronary disease) and any  
XX involving abnormal angiogenesis, neurodegeneration and/or  
XX infectious diseases.  
XX SQ Sequence 2286 BP; 692 A; 422 C; 479 G; 692 T; 1 other;  
  
Query Match 3.2%; Score 33; DB 22; Length 2286;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 664 ttctccattgttaccaggctggtctcgaactc 696  
Db 2111 TTTCGCCATGTTGACCAGGCTGCTCGACTC 2079  
|||||  
RESULT 14  
AAI60415/C  
ID AAI60415 standard; cDNA; 5600 BP.  
XX AC AAI60415;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polynucleotide SEQ ID NO 4404.  
XX KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;  
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
XX KW ankyrotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
XX KW leukaemia; ss.  
XX OS Homo sapiens.  
XX PN W0200153312-A1.  
XX PD 26-JUL-2001.  
XX PF 26-DEC-2000; 2000WO-US34263.  
XX PR 21-JAN-2000; 2000US-0488725.  
XX PR 25-APR-2000; 2000US-0552317.  
XX PR 09-JUL-2000; 2000US-0598042.  
XX PR 19-JUL-2000; 2000US-0620312.  
XX PR 03-AUG-2000; 2000US-0653450.  
XX PR 14-SEP-2000; 2000US-0662191.  
XX PR 19-OCT-2000; 2000US-0693036.  
XX PR 29-NOV-2000; 2000US-0727344.  
XX PA (HYSE-) HYSEQ INC.  
XX PT Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
XX P-PSDB: AAM41259.  
XX DR

XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Claim 1; SEQ ID NO 4404; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 5600 BP; 1708 A; 1046 C; 1174 G; 1672 T; 0 other;

Query Match 3.2%; Score 33; DB 22; Length 5600;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttctgcatgttgaccaggtggtctcgaactc 696  
|||||  
Db 4303 TTTCGCATCTTGACGAGGTGCTCGAATC 4271

RESULT 15  
AAI58629/c  
ID AAI58629 standard; cDNA: 5027 BP.  
XX  
AC AAI58629;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 832.  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
XX  
PR 25-APR-2000; 2000US-0552317.  
XX  
PR 09-JUL-2000; 2000US-0598042.  
XX  
PR 19-JUL-2000; 2000US-0620312.  
XX  
PR 03-AUG-2000; 2000US-0653450.  
XX  
PR 14-SEP-2000; 2000US-0662191.  
XX  
PR 19-OCT-2000; 2000US-0693036.  
XX  
PR 29-NOV-2000; 2000US-0727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI: 2001-442253/47.  
DR P-PSDB; AAM39473.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Claim 1; SEQ ID NO 832; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 6027 BP; 1786 A; 1180 C; 1301 G; 1760 T; 0 other;

Query Match 3.2%; Score 33; DB 22; Length 6027;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 ttctgcatgttgaccaggtggtctcgaactc 696  
|||||  
Db 4643 TTTCGCATCTTGACGAGGTGCTCGAATC 4611

RESULT 16  
AAI64117/c  
ID AAI64117 standard; cDNA: 32191 BP.  
XX  
AC AAI64117;  
XX  
DT 24-OCT-2001 (first entry)  
XX  
DE Human bladder related polynucleotide, SEQ ID NO: 150.  
XX  
KW Human; bladder antigen; cytostatic; immunosuppressive; nootropic;  
KW neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;  
KW antiinflammatory; antitumor; antitumor; antitumor; antitumor;  
KW antifungal; antiparasitic; cardiant; gene therapy; cancer;  
KW immune disorder; cardiovascular disorder; wound healing; infection;  
KW neurological disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200159064-A2.  
XX  
PD 16-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01342.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
XX  
PR 04-FEB-2000; 2000US-0180628.  
XX  
PR 24-FEB-2000; 2000US-0184664.  
XX  
PR 02-MAR-2000; 2000US-0186350.  
XX  
PR 16-MAR-2000; 2000US-0189874.  
XX  
PR 17-MAR-2000; 2000US-0190076.  
XX  
PR 18-APR-2000; 2000US-0198123.  
XX  
PR 19-MAY-2000; 2000US-0205515.  
XX  
PR 07-JUN-2000; 2000US-0209467.



us-09-671-050-9.oli20.rng

Fri May 3 11:02:10 2002

PR	28-JUN-2000;	2000US-0214986.	PR	20-OCT-2000;	2000US-0241786.
PR	30-JUN-2000;	2000US-0215135.	PR	20-OCT-2000;	2000US-0241787.
PR	07-JUL-2000;	2000US-0216647.	PR	20-OCT-2000;	2000US-0241808.
PR	07-JUL-2000;	2000US-0216880.	PR	20-OCT-2000;	2000US-0241809.
PR	11-JUL-2000;	2000US-0217487.	PR	20-OCT-2000;	2000US-0241826.
PR	11-JUL-2000;	2000US-0217496.	PR	01-NOV-2000;	2000US-0244617.
PR	14-JUL-2000;	2000US-0218290.	PR	08-NOV-2000;	2000US-0246474.
PR	26-JUL-2000;	2000US-0220963.	PR	08-NOV-2000;	2000US-0246475.
PR	26-JUL-2000;	2000US-0220964.	PR	08-NOV-2000;	2000US-0246476.
PR	14-AUG-2000;	2000US-0224518.	PR	08-NOV-2000;	2000US-0246477.
PR	14-AUG-2000;	2000US-0224519.	PR	08-NOV-2000;	2000US-0246478.
PR	14-AUG-2000;	2000US-0225213.	PR	08-NOV-2000;	2000US-0246523.
PR	14-AUG-2000;	2000US-0225214.	PR	08-NOV-2000;	2000US-0246524.
PR	14-AUG-2000;	2000US-0225266.	PR	08-NOV-2000;	2000US-0246525.
PR	14-AUG-2000;	2000US-0225267.	PR	08-NOV-2000;	2000US-0246526.
PR	14-AUG-2000;	2000US-0225268.	PR	08-NOV-2000;	2000US-0246527.
PR	14-AUG-2000;	2000US-0225270.	PR	08-NOV-2000;	2000US-0246528.
PR	14-AUG-2000;	2000US-0225270.	PR	08-NOV-2000;	2000US-0246532.
PR	14-AUG-2000;	2000US-0225447.	PR	08-NOV-2000;	2000US-0246609.
PR	14-AUG-2000;	2000US-0225577.	PR	08-NOV-2000;	2000US-0246610.
PR	14-AUG-2000;	2000US-0225758.	PR	08-NOV-2000;	2000US-0246611.
PR	14-AUG-2000;	2000US-0225759.	PR	08-NOV-2000;	2000US-0246613.
PR	18-AUG-2000;	2000US-0226279.	PR	17-NOV-2000;	2000US-0249207.
PR	22-AUG-2000;	2000US-0226681.	PR	17-NOV-2000;	2000US-0249208.
PR	22-AUG-2000;	2000US-0226686.	PR	17-NOV-2000;	2000US-0249209.
PR	22-AUG-2000;	2000US-0227182.	PR	17-NOV-2000;	2000US-0249210.
PR	23-AUG-2000;	2000US-0227009.	PR	17-NOV-2000;	2000US-0249211.
PR	30-AUG-2000;	2000US-0228924.	PR	17-NOV-2000;	2000US-0249212.
PR	01-SEP-2000;	2000US-0229287.	PR	17-NOV-2000;	2000US-0249213.
PR	01-SEP-2000;	2000US-0229343.	PR	17-NOV-2000;	2000US-0249214.
PR	01-SEP-2000;	2000US-0229344.	PR	17-NOV-2000;	2000US-0249215.
PR	01-SEP-2000;	2000US-0229345.	PR	17-NOV-2000;	2000US-0249216.
PR	05-SEP-2000;	2000US-0229509.	PR	17-NOV-2000;	2000US-0249217.
PR	05-SEP-2000;	2000US-0229513.	PR	17-NOV-2000;	2000US-0249218.
PR	06-SEP-2000;	2000US-0230437.	PR	17-NOV-2000;	2000US-0249218.
PR	06-SEP-2000;	2000US-0230438.	PR	17-NOV-2000;	2000US-0249244.
PR	08-SEP-2000;	2000US-0231242.	PR	17-NOV-2000;	2000US-0249245.
PR	08-SEP-2000;	2000US-0231243.	PR	17-NOV-2000;	2000US-0249264.
PR	08-SEP-2000;	2000US-0231244.	PR	17-NOV-2000;	2000US-0249265.
PR	08-SEP-2000;	2000US-0231244.	PR	17-NOV-2000;	2000US-0249265.
PR	08-SEP-2000;	2000US-0231413.	PR	17-NOV-2000;	2000US-0249297.
PR	08-SEP-2000;	2000US-0231414.	PR	17-NOV-2000;	2000US-0249299.
PR	08-SEP-2000;	2000US-0232080.	PR	17-NOV-2000;	2000US-0249300.
PR	08-SEP-2000;	2000US-0232081.	PR	01-DEC-2000;	2000US-0250160.
PR	12-SEP-2000;	2000US-0231968.	PR	01-DEC-2000;	2000US-0250391.
PR	14-SEP-2000;	2000US-0232397.	PR	05-DEC-2000;	2000US-0251030.
PR	14-SEP-2000;	2000US-0232398.	PR	05-DEC-2000;	2000US-0251988.
PR	14-SEP-2000;	2000US-0232399.	PR	05-DEC-2000;	2000US-0256719.
PR	14-SEP-2000;	2000US-0232400.	PR	06-DEC-2000;	2000US-0251479.
PR	14-SEP-2000;	2000US-0232401.	PR	08-DEC-2000;	2000US-0251856.
PR	14-SEP-2000;	2000US-0233063.	PR	08-DEC-2000;	2000US-0251868.
PR	14-SEP-2000;	2000US-0233064.	PR	08-DEC-2000;	2000US-0251869.
PR	14-SEP-2000;	2000US-0233065.	PR	08-DEC-2000;	2000US-0251989.
PR	21-SEP-2000;	2000US-0234223.	PR	08-DEC-2000;	2000US-0251990.
PR	21-SEP-2000;	2000US-0234274.	PR	11-DEC-2000;	2000US-0254097.
PR	25-SEP-2000;	2000US-0234997.	PR	05-JAN-2001;	2001US-0259678.
PR	25-SEP-2000;	2000US-0234998.	XX		
PR	26-SEP-2000;	2000US-0235484.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834.	XX		
PR	27-SEP-2000;	2000US-0235836.	PI	Rosen CA, Barash SC, Ruben SM;	
PR	27-SEP-2000;	2000US-0235836.	XX		
PR	29-SEP-2000;	2000US-0236327.	XX	WPI; 2001-514652/56.	
PR	29-SEP-2000;	2000US-0236367.	XX		
PR	29-SEP-2000;	2000US-0236368.	XX		
PR	29-SEP-2000;	2000US-0236369.	PT	Forty five bladder related polynucleotides, useful in the prevention,	
PR	29-SEP-2000;	2000US-0236370.	PT	treatment and diagnosis of cancer, immune disorders, cardiovascular	
PR	02-OCT-2000;	2000US-0236802.	XX	disorders and neurological diseases -	
PR	02-OCT-2000;	2000US-0237037.	XX	Disclosure; SEQ ID NO 150; 482pp + sequence listing; English.	
PR	02-OCT-2000;	2000US-0237038.	XX		
PR	02-OCT-2000;	2000US-0237039.	XX		
PR	02-OCT-2000;	2000US-0237040.	CC	The invention relates to forty five novel bladder related	
PR	13-OCT-2000;	2000US-0239935.	CC	polynucleotides. The polynucleotides and the polypeptides that they	
PR	13-OCT-2000;	2000US-0239937.	CC	encode are useful in the diagnosis, treatment and prevention of:	
PR	20-OCT-2000;	2000US-0240960.	CC	cancer, particularly breast and ovarian cancer, and other cancers	
PR	20-OCT-2000;	2000US-0241221.	CC	of the adrenal gland, bone, bone marrow, breast, gastrointestinal	
PR	20-OCT-2000;	2000US-0241785.	CC	tract, liver, lung, or urogenital system; immune disorders such as	

CC Addison's disease, allergies, autoimmune haemolytic anaemia,  
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple  
 CC sclerosis, rheumatoid arthritis and ulcerative colitis; cardiovascular  
 CC disorders such as myocardial ischaemia; wound healing; neurological  
 CC diseases such as cerebral anoxia and epilepsy; and infectious diseases  
 CC such as viral, bacterial, fungal and parasitic infections. Numerous  
 CC examples of each type of disorder are given in the specification.  
 CC The polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities. The polynucleotides are  
 CC useful for chromosome identification. They are also useful as probes  
 CC for diagnosing or treating a disorder related to the female reproductive  
 CC system, particularly breast and/or ovary cancer. The present  
 CC sequence is a bladder related polynucleotide encompassed by the  
 CC invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 32191 BP; 10213 A; 6968 C; 6580 G; 8430 T; 0 other;

Query Match 3.0%; Score 31; DB 22; Length 32191;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 660 ggggttcgcacattgaccaggctgctc 690  
 |||||  
 Db 18731 GGGGTTTCGCCATTGACCGAGCTGCTC 18701

RESULT 17  
 AAX98426/c

ID AAX98426 standard; cDNA; 300 BP.

XX AAX98426;

XX 24-SEP-1999 (first entry)

XX Human cancer cell derived cDNA #152.

XX Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;  
 KW integral membrane protein; aspartyl protease; GATA family; wnt family;  
 KW transcription factor; G-protein alpha subunit; protein phosphatase;  
 KW phorbol ester binding protein; diacylglycerol binding protein; trypsin;  
 KW protein kinase; tyrosine phosphatase; developmental signalling protein;  
 KW WW/rsp5/WWP domain; therapy; forensic; genetic mapping; diagnostic;  
 KW detection; treatment; cervical; melanoma; colorectal adenocarcinoma;  
 KW Wilm's tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;  
 KW leukemia; lymphoma; dysplasia; hyperplasia; hyperplasia; endometrium; adrenal;  
 KW prostate; ss.

XX Homo sapiens.

XX W09933982-A2.

XX 08-JUL-1999.

XX 22-DEC-1998; 98WO-US27610.

XX 21-DEC-1998; 98US-0217471.

XX 23-DEC-1997; 97US-0068755.

XX 03-APR-1998; 98US-0080664.

XX 21-OCT-1998; 98US-0105234.

XX 27-OCT-1998; 98US-0105877.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;

XX Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;

XX Jones LH, Kassam A, Kennedy GC, Kita D, Labat I;

XX Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;

XX Stache-Crain B, Sudduth-Klinger J, Williams LT;

XX

DR WPI; 1999-430243/36.

XX New isolated human polynucleotides

XX Claim 1; Page 354; 591pp; English.

XX This invention describes novel isolated human polynucleotides obtained  
 CC by screening for differential expression in colon cancer, breast cancer  
 CC and lung cancer cell lines. The polynucleotides of the invention are  
 CC represented in AAX98275-X99118 and encode polypeptides of protein  
 CC families selected from 4 transmembrane segments integral membrane  
 CC proteins, 7 transmembrane receptors, ATPases associated with various  
 CC cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of  
 CC transcription factors, G-protein alpha subunit, phorbol esters or  
 CC diacylglycerol binding proteins, protein kinase, protein phosphatase 2C,  
 CC protein tyrosine phosphatase, trypsin, wnt family of developmental  
 CC signalling proteins and WW/rsp5/WWP domain containing proteins. The  
 CC encoded polypeptides also have a functional domain selected from Ank  
 CC repeat, basic region plus leucine zipper transcription factors,  
 CC bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger  
 CC (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease  
 CC domain. The polynucleotides encode polypeptides with similarity to known  
 CC protein families and are predicted to have similar properties. The novel  
 CC polynucleotides can be used to develop products for use as therapeutic  
 CC agents and in forensics, genetic analysis, mapping and diagnostic  
 CC applications. In particular, the product can be used for the detection  
 CC and management of cancers. They can be used for treating e.g. cervical  
 CC cancers, melanomas, colorectal adenocarcinomas, Wilm's tumour, sarcomas,  
 CC retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic  
 CC myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and  
 CC hereditary ectodermal dysplasia, congenital alveolar dysplasia, anhydric  
 CC epithelial dysplasia of the cervix, fibrous dysplasia of bone, and  
 CC mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast,  
 CC prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of  
 CC the skin.

SQ Sequence 300 BP; 91 A; 58 C; 66 G; 85 T; 0 other;

Query Match 2.9%; Score 30; DB 20; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-05;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 ggggttcgcacattgaccaggctgctc 689  
 |||||

Db 249 GGGGTTTCGCCATTGACCGAGCTGCTC 220

RESULT 18  
 AAX98945/c

ID AAX98945 standard; cDNA; 752 BP.

XX AAX98945;

XX 24-SEP-1999 (first entry)

XX Human validated cancer cell derived cDNA #267.

XX Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;  
 KW integral membrane protein; aspartyl protease; GATA family; wnt family;  
 KW transcription factor; G-protein alpha subunit; protein phosphatase;  
 KW phorbol ester binding protein; diacylglycerol binding protein; trypsin;  
 KW protein kinase; tyrosine phosphatase; developmental signalling protein;  
 KW WW/rsp5/WWP domain; therapy; forensic; genetic mapping; diagnostic;  
 KW detection; treatment; cervical; melanoma; colorectal adenocarcinoma;  
 KW Wilm's tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;  
 KW leukemia; lymphoma; dysplasia; hyperplasia; hyperplasia; endometrium; adrenal;  
 KW prostate; ss.

XX Homo sapiens.

XX

RESULT 19  
AAC93349  
TID AAC93349 standard: CDNA: 1345 BP.

AC AAH72939;  
 XX  
 DT 19-SEP-2001 (first entry)  
 XX  
 DE Human cervical cancer marker nucleic acid 4213.  
 XX  
 KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200142467-A2.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 08-DEC-2000; 2000WO-US33312.  
 XX  
 PR 08-DEC-1999; 99US-0169681.  
 PR 21-DEC-1999; 99US-0171350.  
 PR 14-MAR-2000; 2000US-0189315.  
 PR 12-MAY-2000; 2000US-0203791.  
 PR 09-JUN-2000; 2000US-0210600.  
 PR 21-JUL-2000; 2000US-0220114.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 XX Schlegel R, Deeds J, Berger A, Zhao X;  
 PI  
 XX WPI; 2001-375006/39.  
 DR  
 XX  
 PT New isolated nucleic acid for diagnosing and treating cervical cancer  
 PT and for assessing and detecting compounds for treating the cancer -  
 XX  
 PS Claim 1; Page 922; 1051pp; English.  
 XX  
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with  
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded  
 CC polypeptides are useful: to assess if a patient is afflicted with  
 CC cervical cancer or has a pre-malignant condition; to monitor the  
 CC progression of cervical cancer or a premalignant condition in a patient;  
 CC and to select and/or assess the efficacy of a compound or therapy for  
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
 CC useful for gene therapy.  
 XX  
 SQ Sequence 1718 BP; 516 A; 407 C; 303 G; 482 T; 10 other;  
 XX  
 Query Match 2.9%; Score 30; DB 22; Length 1718;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 661 ggggttcgccatgttgaccagcgtgtctc 690  
 Db ||||||||||||||||||||||||||||  
 19 ggggttcgccatgttgaccagcgtgtctc 48  
 RESULT 21  
 AAS08146/C  
 ID AAS08146 standard; cDNA; 1762 BP.  
 XX  
 AC AAS08146;  
 XX  
 DT 23-OCT-2001 (first entry)  
 XX  
 DE Mammalian toxicological response marker #35.  
 XX  
 KW Mammalian toxicological response marker; antigen; antibody; agonist; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200136684-A2.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF  
 XX

PF 16-NOV-2000; 2000WO-US31743.  
 XX  
 PR 19-NOV-1999; 99US-0443184.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PT Cunningham MJ, Zweiger GB, Kaser MR, Panzer SR, Seilhamer JJ;  
 PI Yue H, Baughn MR, Azimzai Y, Lal P;  
 XX WPI; 2001-355646/37.  
 DR  
 XX  
 PT Novel mammalian nucleic acid molecules whose levels are up regulated or  
 PT down regulated following treatment with a toxic compound, useful for  
 PT detecting metabolic and toxicological responses and in monitoring drug  
 PT action -  
 XX  
 PS Claim 1; Page 48; 89pp; English.  
 XX  
 CC The sequence is a novel mammalian nucleic acid molecule whose levels are  
 CC up regulated or down regulated following treatment with a toxic compound.  
 CC Polynucleotide sequences complementary to the sequences of the invention  
 CC are useful for preventing a toxicological response by acting against one  
 CC or more up-regulated nucleic acid molecules. An agonist identified by the  
 CC above method is also useful for preventing a toxicological response by  
 CC initiating transcription of a gene comprising a down regulated nucleic  
 CC acid molecule of the invention. Proteins encoded by the nucleic acid of  
 CC the invention are useful for producing antibodies. The nucleic acid of  
 CC molecules are useful for producing an animal model system. The nucleic  
 CC acid immobilised on a substrate as hybridisable array element in a  
 CC microarray format may be used to characterise gene expression patterns  
 CC or to monitor the effects of treatments during clinical trials or therapy  
 CC where metabolic response to toxic compounds may be expected.  
 CC The nucleic acids are useful for various hybridisation technologies and  
 CC is useful for designing hybridisation probes. The nucleic acid molecule  
 CC or its fragment, or a protein encoded by the nucleic acid molecule may be  
 CC used to purify a ligand from a sample.  
 XX  
 SQ Sequence 1762 BP; 497 A; 353 C; 443 G; 469 T; 0 other;  
 XX  
 Query Match 2.9%; Score 30; DB 22; Length 1762;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-05;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 660 ggggttcgccatgttgaccagcgtgtctc 689  
 Db ||||||||||||||||||||||||||||  
 1585 GGGGTTTCGCCATGTTGACCGAGCTGTCT 1556  
 RESULT 22  
 AAH17495  
 ID AAH17495 standard; cDNA; 1804 BP.  
 XX  
 AC AAH17495;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:16962.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.

us-09-671-050-9.oli20.rng

Fri May 3 11:02:10 2002

```
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 16962; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dt primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 1804 BP; 440 A; 453 C; 470 G; 441 T; 0 other;
SQ
Query Match 2.9%; Score 30; DB 22; Length 1804;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 672 tgttgaccaggctggtctcgaaactcttgac 701
Db 1258 tgttgaccaggctggtctcgaaactcttgac 1287
RESULT 23
AAH18014/C
ID AAH18014 standard; cDNA; 1916 BP.
XX
XX AAH18014;
AC
XX
XX 26-JUN-2001 (first entry)
DT
XX
XX Human cDNA sequence SEQ ID NO:17826.
DE
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX Homo sapiens.
OS
XX EP1074617-A2.
PN
XX 07-FEB-2001.
PD
XX 28-JUL-2000; 2000EP-0116126.
PF
XX 29-JUL-1999; 99JP-0248036.
PR
```

```
27-AUG-1999; 99JP-0300253.
11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 17826; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dt primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 1916 BP; 566 A; 352 C; 405 G; 593 T; 0 other;
SQ
Query Match 2.9%; Score 30; DB 22; Length 1916;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 660 ggggtttcgcacatgttgaccaggctggtct 689
Db 1314 GGGGTTTCGCATGTTGACCAGGCTGGTCT 1285
RESULT 24
AAI63723
ID AAI63723 standard; DNA; 9439 BP.
XX
XX AAI63723;
AC
XX
XX 22-OCT-2001 (first entry)
DT
XX
XX Human kidney related polynucleotide SEQ ID NO 1038.
DE
XX
XX Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neutrophic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antiulcer; vulnular; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
XX
XX Homo sapiens.
OS
```

XX	PR	27-SEP-2000;	2000US-0235833.
PN	PR	29-SEP-2000;	2000US-0236327.
XX	PR	29-SEP-2000;	2000US-0236327.
PD	PR	29-SEP-2000;	2000US-0236367.
XX	PR	29-SEP-2000;	2000US-0236369.
PF	PR	23-SEP-2000;	2000US-0236370.
XX	PR	23-SEP-2000;	2000US-0236370.
XX	PR	02-OCT-2000;	2000US-0236802.
PR	PR	02-OCT-2000;	2000US-0237037.
PR	PR	02-OCT-2000;	2000US-0237037.
PR	PR	02-OCT-2000;	2000US-0237038.
PR	PR	02-OCT-2000;	2000US-0237039.
PR	PR	02-OCT-2000;	2000US-0237040.
PR	PR	13-OCT-2000;	2000US-0239935.
PR	PR	13-OCT-2000;	2000US-0239937.
PR	PR	20-OCT-2000;	2000US-0240960.
PR	PR	20-OCT-2000;	2000US-0241221.
PR	PR	20-OCT-2000;	2000US-0241785.
PR	PR	20-OCT-2000;	2000US-0241786.
PR	PR	20-OCT-2000;	2000US-0241787.
PR	PR	20-OCT-2000;	2000US-0241808.
PR	PR	20-OCT-2000;	2000US-0241809.
PR	PR	20-OCT-2000;	2000US-0241826.
PR	PR	01-NOV-2000;	2000US-0244617.
PR	PR	08-NOV-2000;	2000US-0246474.
PR	PR	08-NOV-2000;	2000US-0246475.
PR	PR	08-NOV-2000;	2000US-0246476.
PR	PR	08-NOV-2000;	2000US-0246477.
PR	PR	08-NOV-2000;	2000US-0246478.
PR	PR	08-NOV-2000;	2000US-0246523.
PR	PR	08-NOV-2000;	2000US-0246523.
PR	PR	08-NOV-2000;	2000US-0246524.
PR	PR	08-NOV-2000;	2000US-0246525.
PR	PR	08-NOV-2000;	2000US-0246526.
PR	PR	08-NOV-2000;	2000US-0246527.
PR	PR	08-NOV-2000;	2000US-0246528.
PR	PR	08-NOV-2000;	2000US-0246532.
PR	PR	08-NOV-2000;	2000US-0246609.
PR	PR	08-NOV-2000;	2000US-0246610.
PR	PR	08-NOV-2000;	2000US-0246611.
PR	PR	08-NOV-2000;	2000US-0246613.
PR	PR	17-NOV-2000;	2000US-0249207.
PR	PR	17-NOV-2000;	2000US-0249207.
PR	PR	17-NOV-2000;	2000US-0249208.
PR	PR	17-NOV-2000;	2000US-0249209.
PR	PR	17-NOV-2000;	2000US-0249210.
PR	PR	17-NOV-2000;	2000US-0249211.
PR	PR	17-NOV-2000;	2000US-0249212.
PR	PR	17-NOV-2000;	2000US-0249213.
PR	PR	17-NOV-2000;	2000US-0249214.
PR	PR	17-NOV-2000;	2000US-0249215.
PR	PR	17-NOV-2000;	2000US-0249216.
PR	PR	17-NOV-2000;	2000US-0249217.
PR	PR	17-NOV-2000;	2000US-0249218.
PR	PR	17-NOV-2000;	2000US-0249244.
PR	PR	17-NOV-2000;	2000US-0249245.
PR	PR	17-NOV-2000;	2000US-0249245.
PR	PR	17-NOV-2000;	2000US-0249264.
PR	PR	17-NOV-2000;	2000US-0249265.
PR	PR	17-NOV-2000;	2000US-0249266.
PR	PR	17-NOV-2000;	2000US-0249297.
PR	PR	17-NOV-2000;	2000US-0249299.
PR	PR	01-DEC-2000;	2000US-0249300.
PR	PR	01-DEC-2000;	2000US-0250160.
PR	PR	01-DEC-2000;	2000US-0250391.
PR	PR	05-DEC-2000;	2000US-0251030.
PR	PR	05-DEC-2000;	2000US-0251988.
PR	PR	05-DEC-2000;	2000US-0256719.
PR	PR	06-DEC-2000;	2000US-0251479.
PR	PR	08-DEC-2000;	2000US-0251856.
PR	PR	08-DEC-2000;	2000US-0251868.
PR	PR	08-DEC-2000;	2000US-0251869.
PR	PR	08-DEC-2000;	2000US-0251889.
PR	PR	08-DEC-2000;	2000US-0251990.
PR	PR	11-DEC-2000;	2000US-0254097.
PR	PR	05-JAN-2001;	2001US-0259678.
XX	XX	(HUMA-) HUMAN GENOME SCI INC.	
XX	XX		

PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-488784/53.  
 DR  
 XX New isolated nucleic acids and polypeptides, useful for diagnosing,  
 XX treating and/or preventing human diseases and disorders -  
 PT  
 XX  
 XX Disclosure; SEQ ID NO 1038; 564pp + Sequence Listing; English.  
 PS  
 XX  
 XX The invention relates to novel kidney related polynucleotides  
 CC (AA162971-AA163793) and the encoded polypeptides (AA42417-AA42691)  
 CC collectively known as kidney antigens and the use of such kidney antigens  
 CC for detecting disorders of the kidney, especially kidney cancer and  
 CC kidney cancer metastases. The polynucleotides and proteins are also  
 CC useful for preventing, treating or ameliorating medical conditions  
 CC e.g. by protein or gene therapy. The genes are isolated from a range  
 CC of human tissues disclosed in the specification. The nucleic acids,  
 CC proteins, antibodies and (ant)agonists are useful in the diagnosis,  
 CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,  
 CC and other cancers of the adrenal gland, bone, bone marrow, breast,  
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders  
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,  
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple  
 CC sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound  
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
 CC and (f) infectious diseases such as viral, bacterial, fungal and  
 CC parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 9439 BP; 2744 A; 1972 C; 2076 G; 2647 T; 0 other;  
 SQ

Query Match 2.8%; Score 29; DB 22; Length 9439;  
 Best Local Similarity 100.0%; Pred. No. 0.00015;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 668 gccatgtgaccaggctggtctgaactc 696  
 Db 3769 gccatgtgaccaggctggtctgaactc 3797  
 |||||||||||||||||||||||||||||  
 RESULT 26  
 AAV57903/C  
 ID AAV57903 standard; DNA; 237326 BP.  
 XX  
 XX AAV57903;  
 AC  
 XX  
 XX 21-DEC-1998 (first entry)  
 DT  
 XX  
 XX Hereditary haemochromatosis subregion from an HH affected individual.  
 DE  
 XX  
 XX Bovine butyrophilin; BT: human hereditary haemochromatosis; HFE;  
 XX diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF2; BTF3;  
 XX BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;  
 XX type 1 sodium transport gene; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO9814466-A1.  
 PN  
 XX 09-APR-1998.  
 PD  
 XX 30-SEP-1997; 97WO-US17658.  
 PF  
 XX 07-MAY-1997; 97US-0852495.  
 PR  
 XX 01-OCT-1996; 96US-0724394.  
 PR  
 XX (PROG-) PROGENITOR INC.  
 PA  
 XX Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;  
 XX

PI Tsuchihashi Z, Wolff RK;  
 XX WPI; 1998-240014/21.  
 DR  
 XX Hereditary haemochromatosis gene products - used to develop products  
 XX for the diagnosis and treatment of hereditary disorders in iron  
 PT metabolism  
 PT  
 XX  
 XX Claim 1; Fig 9; 209pp; English.  
 PS  
 XX  
 XX The present invention describes hereditary haemochromatosis gene  
 CC products from the human haemochromatosis gene. The present sequence  
 CC represents a hereditary haemochromatosis subregion from an hereditary  
 CC haemochromatosis (HH) affected individual. Also described is a  
 CC method to determine the presence or absence of the common hereditary  
 CC haemochromatosis (HFE) gene mutation in an individual comprising:  
 CC (a) providing DNA or RNA from the individual; and (b) assessing the  
 CC DNA or RNA for the presence or absence of a haplotype or genotype where  
 CC the presence or absence of the haplotype genotype indicates the likely  
 CC presence of the HFE gene mutation in the genome of the individual. The  
 CC HFE gene sequences from the present invention can be used to develop  
 CC products for use in the diagnosis and treatment of HFE. The present  
 CC invention also describes BTF genes, which are homologues of the milk  
 CC protein butyrophilin (BT), and can be used in the production of agonists  
 CC and antagonists of BT function. Also described are: (1) a RoRet gene  
 CC which can be used to develop products for the study, diagnosis and  
 CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes  
 CC which are homologues of a type 1 sodium transport gene, and can  
 CC similarly be used for hypophosphatemia.  
 XX  
 XX Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;  
 SQ

Query Match 2.8%; Score 29; DB 19; Length 237326;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 662 ggtttcccatgttgaccaggctggtctc 690  
 Db 138329 GGTTCGCATGTTGACCGCTGGCTC 138301  
 |||||||||||||||||||||||||||||  
 RESULT 26  
 AAC05941  
 ID AAC05941 standard; cDNA; 261 BP.  
 XX  
 XX AAC05941;  
 AC  
 XX  
 XX 06-OCT-2000 (first entry)  
 DT  
 XX  
 XX Human secreted protein 5' EST, SEQ ID NO: 10016.  
 DE  
 XX  
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 XX gene therapy; chromosome mapping; ss.  
 KW  
 XX Homo sapiens.  
 OS  
 XX EP1033401-A2.  
 PN  
 XX 06-SEP-2000.  
 PD  
 XX 21-FEB-2000; 2000EP-0200610.  
 PF  
 XX 26-FEB-1999; 99US-0122487.  
 PR  
 XX (GEST ) GENSET.  
 PA  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 XX WPI; 2000-500381/45.  
 DR  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX Claim 1; SEQ ID 10016; 71pp + CD-ROM; English.  
 XX The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
 CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors.  
 XX  
 SQ Sequence 261 BP; 62 A; 52 C; 55 G; 90 T; 2 other;

Query Match 2.7%; Score 28; DB 21; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 0.00044;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 650 tagtagagacgggttcgcacatgttga 677  
 |||||  
 DB 13 tagtagagacgggttcgcacatgttga 40

RESULT 27  
 AAC05857  
 ID AAC05857 standard; cDNA; 323 BP.  
 XX  
 AC AAC05857;  
 XX  
 DT 06-OCT-2000 (first entry)  
 XX  
 DE Human secreted protein 5' EST, SEQ ID NO: 9932.  
 XX  
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 XX Homo sapiens.  
 XX EP1033401-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 21-FEB-2000; 2000EP-0200610.  
 XX  
 PR 26-FEB-1999; 99US-0122487.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Dumas Milne Edwards J, Duclert A, Giordano J;  
 XX WPI; 2000-500381/45.  
 XX  
 PS New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 PS Claim 1; SEQ ID 9932; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences

CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors.  
 XX  
 SQ Sequence 323 BP; 46 A; 86 C; 80 G; 110 T; 1 other;

Query Match 2.7%; Score 28; DB 21; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 0.00045;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 669 ccattgtgaccaggctggtctcgaactc 696  
 |||||  
 DB 248 ccattgtgaccaggctggtctcgaactc 275

RESULT 28  
 AAHL3269  
 ID AAHL3269 standard; cDNA; 531 BP.  
 XX  
 AC AAHL3269;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA clone (3'-primer) SEQ ID NO:10104.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX Homo sapiens.  
 XX EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 3; SEQ ID 10104; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and



CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX SQ Sequence 531 BP; 112 A; 147 C; 121 G; 144 T; 7 other;

Query Match 2.7%; Score 28; DB 22; Length 531;  
Best Local Similarity 100.0%; Pred. No. 0.00045;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccaggctggtctcgaaactc 696  
Db 175 ccattgtgaccaggctggtctcgaaactc 202  
|||||

RESULT 29  
AAH16349  
ID AAA16349 standard; DNA; 631 BP.  
XX  
AC AAA16349;  
XX  
DT 14-JUN-2000 (first entry)  
XX  
DE Human colon cancer differentially expressed nucleotide sequence #354.  
XX  
KW Colon cancer; detect; differential expression; human; treatment;  
KW detect mutation; non-invasive diagnostic method; ds.  
XX  
OS Homo sapiens.  
XX  
PN WQ200012702-A2.  
XX  
PD 09-MAR-2000.  
XX  
PF 30-AUG-1999; 99WO-US19424.  
XX  
PR 31-AUG-1998; 98US-0098639.  
XX 27-JAN-1999; 99US-0117393.  
XX  
PA (FARB ) BAYER CORP.  
PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Carroll E;  
PI Catino TJ, Dwiwedi P, Ford DM, Lewis ME, Molino GA, Monahan JE;  
PI Schlegel R;  
XX  
XX WPI; 2000-256641/22.  
XX  
XX Novel nucleic acids and proteins for identifying therapeutic agents  
PT useful for treating and diagnosing cancer, especially colon cancer -  
PT  
XX  
XX Claim 16; Page 267; 345pp; English.

CC This sequence represents a human nucleotide sequence which is  
CC differentially expressed in colon cancer cells compared to the expression  
CC levels in normal cells. The nucleotide sequence can be used as a source  
CC of primers and probes. The nucleotide sequence is useful for determining  
CC the phenotype of a cell by detecting the differential expression of the  
CC sequence relative to a normal cell. The probes derived from the sequence  
CC can also be used to determine the phenotype of cells in a sample. Probes  
CC and antibodies which hybridise to the nucleotide sequence can also be  
CC used to determine the phenotype of a cell. The primers are useful for  
CC detecting a mutation in a test nucleotide sequence and also for detecting  
CC cancer, preferably colon cancer. Antibodies against the protein encoded  
CC by the nucleotide sequence can also be used in a method to detect colon  
CC cancer. The diagnostic method is non-invasive and accurate for diagnosing  
CC colon cancer at an early stage.

XX SQ Sequence 631 BP; 124 A; 162 C; 154 G; 171 T; 20 other;

Query Match 2.7%; Score 28; DB 21; Length 631;  
Best Local Similarity 100.0%; Pred. No. 0.00045;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccaggctggtctcgaaactc 696  
Db 214 ccattgtgaccaggctggtctcgaaactc 241  
|||||

RESULT 30  
AAH04221  
ID AAH04221 standard; cDNA; 805 BP.  
XX  
AC AAH04221;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human cDNA clone (5'-primer) SEQ ID NO:1056.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.  
XX  
PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
XX WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX  
XX Claim 1; SEQ ID 1056; 2537pp + CD ROM; English.

CC The present invention describes primer sets for synthesising 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC

CC represent oligonucleotides, all of which are used in the exemplification  
 XX of the present invention.

SQ Sequence 805 BP; 203 A; 200 C; 188 G; 207 T; 7 other;

Query Match 2.7%; Score 28; DB 22; Length 805;  
 Best Local Similarity 100.0%; Pred. No. 0.00045;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 663 gtttcgcacattgacacagtgctc 690  
 |||||

Db 200 gtttcgcacattgacacagtgctc 227

RESULT 31

AAH51279/c

ID AAH51279 standard; DNA; 1001 BP.

XX

AC AAH51279;

XX

DT 29-AUG-2001 (first entry)

XX

DE Human GSHS related DNA containing a biallelic polymorphism SEQ ID 170.

XX

KW Human; biallelic marker; single nucleotide polymorphism; SNP; MGSTII;  
 KW microsomal glutathione S-transferase II; malate decarboxylase enzyme;  
 KW DME1; ME1; cytochrome P450; glutathione reductase; GSHR; GSHS; GGT5;  
 KW flavin-containing monooxygenase; FMO; gamma-glutamyltransferase 5;  
 KW dipeptidase; DP; glucose 6-phosphate dehydrogenase; G6PDH; haplotype;  
 KW phosphogluconate dehydrogenase; PGDH; drug metabolism; phenotype;  
 KW uridine diphosphate glucuronosyl transferase; UGT2; asthma; hepatotoxicity;  
 KW zileuton; ds.

XX

OS Homo sapiens.

XX

PN WO200058508-A2.

XX

PD 05-OCT-2000.

XX

PF 24-MAR-2000; 2000WO-IB00403.

XX

PR 25-MAR-1999; 99US-0126269.

XX

PR 30-APR-1999; 99US-0131961.

XX

PA (GEST ) GENSET.

XX

PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen-Akenine A;

XX

DR WPI; 2000-638353/61.

XX

PT Polynucleotides comprising sequences from malate decarboxylase

XX

PS enzyme-related biallelic markers used for genotyping -

XX

Claim 13; Page 375-376; 673pp; English.

XX

CC Sequences AAH51110-AAH51593 represent human DNA fragments which contain  
 CC biallelic markers. The sequences are related to various human genes  
 CC including microsomal glutathione S-transferase II (MGSTII), malate  
 CC decarboxylase enzyme (DME1/ME1), cytochrome P450, glutathione  
 CC reductase/synthase (GSHR/GSHS), flavin-containing monooxygenases (FMO),  
 CC gamma-glutamyltransferase 5 (GGT5), dipeptidase (DP), glucose 6-phosphate  
 CC dehydrogenase (G6PDH), phosphogluconate dehydrogenase (PGDH) and uridine  
 CC diphosphate glucuronosyl transferases (UGT2). Each of these sequences  
 CC contains a biallelic marker/polymorphism which is represented in the  
 CC sequence as a degenerate/undefined base. The genes to which the biallelic  
 CC marker containing sequences are related are involved in drug metabolism.  
 CC Sequences AAH51594 - AAH51598 represent the genomic sequence of the  
 CC MGSTII gene and four alternative MGSTII cDNA sequences. AAB62905-AAH62906  
 CC are MGSTII gene products. PCR primers AAH51599 and AAH51600 are used in  
 CC an example for the amplification of human genomic DNA fragments. The  
 CC invention includes a method of genotyping comprising determining the  
 CC identity of a nucleotide at a DME- or MGSTII-related biallelic marker in

CC a biological sample. The method is used to determine the frequency in  
 CC population of an allele of a DME- or MGST-II related biallelic marker and  
 CC to select an individual for inclusion in a clinical trial of a drug  
 CC treatment. The method is also used to detect association between allele  
 CC and phenotype, and to detect association between haplotype and phenotype.  
 CC The polynucleotides are used, in hybridization assays, sequencing assays  
 CC or allele specific amplification assays. The method can be used to  
 CC determine whether an individual suffers or is at risk of developing  
 CC asthma or is at risk of developing hepatotoxicity on treatment with  
 CC zileuton.

XX  
 SQ Sequence 1001 BP; 340 A; 200 C; 211 G; 249 T; 1 other;

Query Match 2.7%; Score 28; DB 21; Length 1001;

Best Local Similarity 100.0%; Pred. No. 0.00045;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ccattgtgaccagctgctcgaactc 696  
 |||||

Db 745 CCATGTTGACGAGGCTGCTCGAATC 718

RESULT 32

AAH51559/c

ID AAH51559 standard; DNA; 1001 BP.

XX

AC AAH51559;

XX

DT 29-AUG-2001 (first entry)

XX

DE Human GLCL related DNA containing a biallelic polymorphism SEQ ID 450.

XX

KW Human; biallelic marker; single nucleotide polymorphism; SNP; MGSTII;  
 KW microsomal glutathione S-transferase II; malate decarboxylase enzyme;  
 KW DME1; ME1; cytochrome P450; glutathione reductase; GSHR; GSHS; GGT5;  
 KW flavin-containing monooxygenase; FMO; gamma-glutamyltransferase 5;  
 KW dipeptidase; DP; glucose 6-phosphate dehydrogenase; G6PDH; haplotype;  
 KW phosphogluconate dehydrogenase; PGDH; drug metabolism; phenotype;  
 KW uridine diphosphate glucuronosyl transferase; UGT2; asthma; hepatotoxicity;  
 KW zileuton; ds.

XX

OS Homo sapiens.

XX

PN WO200058508-A2.

XX

PD 05-OCT-2000.

XX

PF 24-MAR-2000; 2000WO-IB00403.

XX

PR 25-MAR-1999; 99US-0126269.

XX

PR 30-APR-1999; 99US-0131961.

XX

PA (GEST ) GENSET.

XX

PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen-Akenine A;

XX

DR WPI; 2000-638353/61.

XX

PT Polynucleotides comprising sequences from malate decarboxylase

XX

PS enzyme-related biallelic markers used for genotyping -

XX

Claim 13; Page 624-625; 673pp; English.

XX

CC Sequences AAH51110-AAH51593 represent human DNA fragments which contain  
 CC biallelic markers. The sequences are related to various human genes  
 CC including microsomal glutathione S-transferase II (MGSTII), malate  
 CC decarboxylase enzyme (DME1/ME1), cytochrome P450, glutathione  
 CC reductase/synthase (GSHR/GSHS), flavin-containing monooxygenases (FMO),  
 CC gamma-glutamyltransferase 5 (GGT5), dipeptidase (DP), glucose 6-phosphate  
 CC dehydrogenase (G6PDH), phosphogluconate dehydrogenase (PGDH) and uridine  
 CC diphosphate glucuronosyl transferases (UGT2). Each of these sequences  
 CC contains a biallelic marker/polymorphism, which is represented in the



CC gastrointestinal, developmental and cell proliferative disorders. The  
CC HTPN's can be used to treat or prevent disorders associated with a  
CC decreased expression or activity of HTPN.  
XX  
SQ Sequence 1562 BP; 324 A; 464 C; 404 G; 370 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 1562;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccattgtgaccaggctggtctcgaaactc 696  
Db 1459 ccattgtgaccaggctggtctcgaaactc 1486  
|||||

RESULT 35  
AAH98793  
ID AAH98793 standard; cDNA; 1624 BP.  
XX  
AC AAH98793;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Human EST-derived coding sequence SEQ ID NO: 650.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200154477-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 25-JAN-2001; 2001WO-US02687.  
XX  
PR 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX  
PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Wehrman I;  
XX  
DR WPI: 2001-476164/51.  
DR P-PSDB: AAM24134.  
XX  
PT Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
XX  
PS Claim 1; Page 615; 1275pp; English.  
XX  
CC The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
CC of the invention.  
XX  
SQ Sequence 1624 BP; 426 A; 347 C; 343 G; 507 T; 1 other;

Query Match 2.7%; Score 28; DB 22; Length 1624;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccattgtgaccaggctggtctcgaaactc 696  
Db 312 ccattgtgaccaggctggtctcgaaactc 339  
|||||

RESULT 36  
AAF90192  
ID AAF90192 standard; cDNA; 1763 BP.  
XX  
AC AAF90192;  
XX  
DT 06-AUG-2001 (first entry)  
XX  
DE Nucleotide sequence of human spleen extracted chemokine (SECKine).

XX G protein-coupled receptor ligand; spleen extracted chemokine; SECKine;  
KW bonzo; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 15..779  
FT /tag= a  
FT /product= "SECKine"  
FT /transl\_except= "(pos: 321..323, aa: Glu)"  
FT /transl\_except= "(pos: 381..383, aa: Thr)"  
FT sig\_peptide 15..101  
FT /tag= c  
XX  
PN WO200138874-A1.  
XX  
PD 31-MAY-2001.  
XX  
PF 22-NOV-2000; 2000WO-US32095.  
XX  
PR 24-NOV-1999; 99US-0448725.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Briskin MJ;  
XX  
DR WPI: 2001-367732/38.  
DR P-PSDB: AAB84199.  
XX

PT Isolating genes encoding G-protein-coupled receptor ligand, comprises  
PT expressing an exogenous nucleic acid library in a primary cell pool,  
PT selecting and growing cells that encode ligand, and combining them into  
PT a secondary cell pool -  
XX  
PS Example; Fig 2; 43pp; English.  
XX  
CC The specification describes a method for isolating nucleic acid  
CC encoding a G protein-coupled receptor (GPCR) ligand. The method comprises  
CC expressing an expression library having exogenous nucleic acids  
CC inserted into a primary pool of prokaryotic cells, to isolate a pool  
CC of proteins comprising the ligand, selecting and growing cells that  
CC encode the pool of proteins to produce individual colonies, and  
CC combining them into a secondary pool of cells for recovering the  
CC nucleic acid. The method is useful for isolating a nucleic acid,  
CC preferably a cDNA encoding a ligand for a GPCR, such as a chemokine  
CC receptor. The present sequence encodes a spleen extracted chemokine  
CC (SECKine), a ligand for bonzo. SECKine was isolated using the method  
CC of the invention.  
XX  
SQ Sequence 1763 BP; 374 A; 514 C; 395 G; 480 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 1763;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccattgtgaccaggctggtctcgaaactc 696  
|||||

Db	1256	ccatgttgaccaggctggtctcgaactc	1283	XX	26-JUN-2001	(first entry)	
RESULT	37			DT	Human cDNA sequence	SEQ ID NO:17990.	
AAF85365				DE	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
ID	AAF85365	standard; cDNA; 1763 BP.		KW	Homo sapiens.		
AC	AAF85365;			OS	EP1074617-A2.		
XX	23-JUL-2001	(first entry)		PN	07-FEB-2001.		
DT				XX	28-JUL-2000; 2000EP-0116126.		
DE				XX	29-JUL-1999; 99JP-0248036.		
XX				XX	27-AUG-1999; 99JP-0300253.		
KW				PR	11-JAN-2000; 2000JP-0118776.		
XX				PR	02-MAY-2000; 2000JP-0183767.		
OS				PR	09-JUN-2000; 2000JP-0241899.		
XX				XX	(HELI-) HELIX RES INST.		
XX				PA	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
FT				PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;		
FT				PI	WPT; 2001-318749/34.		
FT				XX	Primer sets for synthesizing polynucleotides, particularly the 5602		
FT				PT	full-length cDNAs defined in the specification, and for the detection		
FT				PT	and/or diagnosis of the abnormality of the proteins encoded by the		
XX				PT	full-length cDNAs -		
XX				PS	Claim 8; SEQ ID 17990; 2537pp + CD ROM; English.		
XX				XX	The present invention describes primer sets for synthesising 5602		
XX				CC	full-length cDNAs defined in the specification. Where a primer set		
XX				CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary		
XX				CC	to the complementary strand of a polynucleotide which comprises one of		
XX				CC	the 5602 nucleotide sequences defined in the specification, where the		
XX				CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination		
XX				CC	of an oligonucleotide comprising a sequence complementary to the		
XX				CC	complementary strand of a polynucleotide which comprises a 5'-end		
XX				CC	sequence and an oligonucleotide comprising a sequence complementary to a		
XX				CC	polynucleotide which comprises a 3'-end sequence, where the		
XX				CC	oligonucleotide comprises at least 15 nucleotides and the combination of		
XX				CC	the 5'-end sequence/3'-end sequence is selected from those defined in		
XX				CC	the specification. The primer sets can be used in antisense therapy and		
XX				CC	in gene therapy. The primers are useful for synthesising polynucleotides,		
XX				CC	particularly full-length cDNAs. The primers are also useful for the		
XX				CC	detection and/or diagnosis of the abnormality of the proteins encoded by		
XX				CC	the full-length cDNAs. The primers allow obtaining of the full-length		
XX				CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and		
XX				CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to		
XX				CC	AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632		
XX				CC	represent oligonucleotides, all of which are used in the exemplification		
XX				CC	of the present invention.		
XX				XX	Sequence 1791 BP; 484 A; 402 C; 433 G; 472 T; 0 other;		
XX				XX	Query Match 2.7%; Score 28; DB 22; Length 1791;		
XX				XX	Best Local Similarity 100.0%; Pred. No. 0.00046;		
XX				XX	Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	669	ccatgttgaccaggctggtctcgaactc	696	QY	669	ccatgttgaccaggctggtctcgaactc	696
Db	1256	ccatgttgaccaggctggtctcgaactc	1283	Db	1767	CCATGTTGACCAAGCTGCTCGAACTC	1740
RESULT	38						
AAH18124/C				RESULT	39		
ID	AAH18124	standard; cDNA; 1791 BP.		AAC59834			
XX				ID	AAC59834	standard; DNA; 1792 BP.	
AC	AAH18124;						

AAC59834;  
 26-JAN-2001 (first entry)  
 Human secreted protein encoding DNA clone vq19 1.  
 Secreted protein; human; autoimmune disorder; multiple sclerosis; ulcer;  
 systemic lupus erythematosus; rheumatoid arthritis; anaemia; stroke;  
 haematopoiesis regulation; tissue regrowth; wound healing; haemophilia;  
 Alzheimer's disease; Parkinson's disease; Shy-drager syndrome; cancer;  
 contraceptive; infection; growth inhibition; hyperproliferative disorder;  
 psoriasis; ds.  
 Homo sapiens.  
 WO200055375-A1.  
 21-SEP-2000.  
 17-MAR-2000; 2000WO-US07285.  
 17-MAR-1999; 99US-0124808.  
 17-MAR-1999; 99US-0124916.  
 17-AUG-1999; 99US-0149639.  
 01-OCT-1999; 99US-0157247.  
 29-NOV-1999; 99US-0167824.  
 15-FEB-2000; 2000US-0182711.  
 (ALPH-) ALPHAGENE INC.  
 Valenzuela D, Yuan O, Hoffman H, Hall J, Raple[j]ko P;  
 WPI; 2000-638211/61.  
 P-PSDB; AAB34733.  
 Novel proteins and polypeptides useful for the treatment of e.g  
 multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,  
 cancer, Alzheimer's disease, Parkinson's disease, stroke, anemia and  
 ulcers -  
 Claim 102; Page 446-447; 493pp; English.  
 This invention relates to 59 human secreted proteins and the nucleotide  
 sequences encoding them. Sequences AAC59788-C59846 and AAB34687-B34745  
 represent the proteins and their encoding nucleotide sequences, and  
 sequences AAB34746-B34771 represent fragments of the proteins. Probes  
 for the DNA sequences are represented by sequences AAC59847-C59596. The  
 proteins exhibit neuroprotective, dermatological, immunosuppressive,  
 antiinflammatory, antianemic, nootropic, antiparkinsonian,  
 cerebroprotective, haemostatic, vulnery, cytostatic, antipsoriatic,  
 antibacterial, virucide, and fungicide activity. The proteins and  
 nucleotide sequences are useful as nutritional sources or supplements  
 and in research. The proteins are useful for treating immune deficiency  
 and disorders, which may be genetic or resulting from immune deficiency.  
 autoimmune disorders such as multiple sclerosis, systemic lupus  
 erythematosus, rheumatoid arthritis, and for treating myeloid or lymphoid  
 cell deficiencies such as anaemias by regulating haematopoiesis. The  
 proteins are also useful in compositions for bone, cartilage, tendon,  
 ligament and/or nerve tissue growth or regeneration, for wound healing,  
 tissue repair and replacement and in the treatment of wounds, incisions  
 and ulcers. Other uses include in the treatment of central and  
 peripheral nervous system and neuropathies such as Alzheimer's and  
 Parkinson's diseases and Shy-Drager syndrome, and mechanical and  
 traumatic disorders, such as spinal cord disorders, head trauma and  
 stroke. The proteins may also be used as a contraceptive, and for  
 treating coagulation disorders such as haemophilias. The protein and  
 nucleotide sequences with cadherin activity are useful for treating  
 cancer. Other uses for the protein include for inhibiting the growth,  
 infection or function of, or killing, infectious agents such as bacteria,  
 virus, fungi and other parasites, for effecting bodily characteristics  
 such as height, weight, hair colour, effecting biorhythms or cardiac  
 cycles or rhythms, effecting metabolism, catabolism, anabolism,

CC processing, utilization, storage or elimination of dietary fat, lipid,  
 CC protein, carbohydrate, vitamins, minerals, cofactors, effecting  
 CC behavioural characteristics, providing analgesic effects and for treating  
 CC hyperproliferative disorders such as psoriasis.  
 XX  
 SQ Sequence 1792 BP; 417 A; 507 C; 387 G; 481 T; 0 other;  
 Query Match 2.7%; Score 28; DB 21; Length 1792;  
 Best Local Similarity 100.0%; Pred. No. 0.00046;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 669 ccattgtgaccaggctgctcgaaactc 696  
 Db 1241 ccattgtgaccaggctgctcgaaactc 1268  
 RESULT 40  
 AAH18088  
 ID AAH18088 standard; cDNA; 1850 BP.  
 XX  
 AC AAH18088;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:17938.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EPI074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 17938; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 1850 BP; 473 A; 444 C; 390 G; 543 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 1850;  
 Best Local Similarity 100.0%; Pred. No. 0.00046;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccattgtgaccaggctgctcgaaactc 696  
 |||||  
 Db 1266 ccattgtgaccaggctgctcgaaactc 1293

RESULT .41  
 AAC76936  
 ID AAC76936 standard; cDNA; 1902 BP.  
 XX  
 AC AAC76936;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF2491 polynucleotide sequence SEQ ID NO:4981.  
 XX  
 KW Human; open reading frame; ORFX: detection; cytostatic; hepatotropic;  
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US08621.  
 XX  
 PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Leach M;  
 XX  
 XX WPI: 2000-602362/57.  
 DR P-PSDB; AAB42727.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 5; Page 4155-4157; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;

CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
 CC antithyroid; and antianaemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 XX  
 SQ Sequence 1902 BP; 366 A; 541 C; 446 G; 548 T; 1 other;

Query Match 2.7%; Score 28; DB 21; Length 1902;  
 Best Local Similarity 100.0%; Pred. No. 0.00046;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccattgtgaccaggctgctcgaaactc 696  
 |||||  
 Db 652 ccattgtgaccaggctgctcgaaactc 679

RESULT 42  
 AAH17566  
 ID AAH17566 standard; cDNA; 2252 BP.  
 XX  
 AC AAH17566;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:17050.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 KW Homo sapiens.  
 OS  
 PN BP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI: 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 17050; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set

comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 2252 BP; 435 A; 692 C; 571 G; 554 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 2252;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 669 ccattgtgaccagcgtgtctcgaactc 696  
||||| ||||||| ||||||| ||||||| |||||||  
Db 1747 ccattgtgaccagcgtgtctcgaactc 1774

## RESULT 43

AAZ25332  
ID AAZ25332 standard; cDNA; 2309 BP.

AC AAZ25332;

DT 20-DEC-1999 (first entry)

DE Human chemokine alpha-5 nucleotide sequence.

KW Human; chemokine alpha-5; Ckalpha-5; diagnosis; detection; immunity; immune system related disorder; solid tumour; angiogenesis; lymphoma; infection; autoimmune disease; lymphocytic leukaemias; wound healing; fibrotic disorder; liver cirrhosis; osteoarthritis; pulmonary fibrosis; haematopoiesis; regulation; autoimmunity; arthritis; leukaemias; immunosuppression; inflammatory bowel disease; myelo suppression; ss.

OS Homo sapiens.

PH Key Location/Qualifiers  
FT CDS 542..1306  
FT sig\_peptide /tag= a  
FT mat\_peptide 542..622  
FT /tag= b  
FT /tag= c

PN WO9927078-A1.

PD 03-JUN-1999.

PF 18-NOV-1998; 98WO-US24619.

PR 21-NOV-1997; 97US-0066369.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Wei Y, Ni J, Li Y, Rosen CA, Ruben SM;

DR WPI; 1999-590675/50.

DR P-PSDB; AAY42165.

PT New isolated chemokine alpha-5

PS Claim 2; Fig 1; 86pp; English.

XX The present sequence encodes human chemokine alpha-5 (Ckalpha-5).  
CC The novel Ckalpha-5 polypeptides have chemokine activity. It can  
CC be used to treat solid tumours, angiogenesis, chronic and acute  
CC infections, e.g. mycobacterial infections, autoimmune diseases and  
CC lymphocytic leukaemias, stimulate wound healing, fibrotic disorders,  
CC (e.g. liver cirrhosis, osteoarthritis and pulmonary fibrosis), killing  
CC the larvae of parasites that invade tissues (as in schistosomiasis,  
CC trichinosis and ascariasis), regulate haematopoiesis, re-vascularisation  
CC of damaged limbs (from injury or disease), dysregulation of immune cell  
CC function (e.g. autoimmunity, arthritis, leukaemias, lymphomas,  
CC immunosuppression, immunity, humoral immunity, inflammatory bowel  
CC disease, or myelo suppression). Antagonists to Ckalpha-5 can be used in  
CC the treatment of autoimmune and chronic inflammatory and infective diseases  
CC (e.g. multiple sclerosis, sarcoidosis, and idiopathic pulmonary  
CC fibrosis) (e.g. silicosis, sarcoidosis, and idiopathic pulmonary  
CC atherosclerosis, histamine-mediated allergic reactions and immunological  
CC disorders (e.g. late phase allergic reactions, chronic urticaria, and  
CC atopic dermatitis), IgE-mediated allergic reactions (e.g. allergic  
CC asthma, rhinitis and eczema) to treat chronic and acute inflammation),  
CC rheumatoid arthritis, bone marrow failure, (e.g. aplastic anaemia and  
CC myelodysplastic syndrome), and subepithelial basement membrane  
CC fibrosis. The products can also be used for detection, diagnosis and  
CC drug screening.

XX Sequence 2309 BP; 462 A; 706 C; 581 G; 560 T; 0 other;

Query Match 2.7%; Score 28; DB 20; Length 2309;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 669 ccattgtgaccagcgtgtctcgaactc 696  
||||| ||||||| ||||||| ||||||| |||||||  
Db 1782 ccattgtgaccagcgtgtctcgaactc 1809

## RESULT 44

AAAY1606

ID AAA71606 standard; DNA; 2575 BP.

AC AAA71606;

DT 11-DEC-2000 (first entry)

DE Human CXC family chemokine-like protein encoding DNA SEQ ID NO: 1.

KW Human; CXC family; chemokine-like; disease; treatment; CL protein; ds.

OS Homo sapiens.

PN WO200044784-A1.

PD 03-AUG-2000.

PF 21-JAN-2000; 2000WO-IL00043.

PR 27-JAN-1999; 99IL-0128252.

PA (COMP-) COMPUGEN LTD.

PI Mintz L, Savitzky K;

XX WPI; 2000-491238/43.

DR P-PSDB; AAB10491.

XX New human chemokine like (CL) nucleic acids and proteins, useful for



PT diagnosing and treating diseases resulting from inherited defective CL  
PT sequences -  
PS Claim 1(i); Page 47-48; 55pp; English.  
XX This invention describes novel human chemokine like (CL) nucleic acids  
XX (N1), their complementary sequences (N2) and encoded amino acids (P1).  
CC An expression vector or P1 is useful for treating diseases which  
CC can be ameliorated or cured by raising the level of the CL protein. N2,  
CC the expression vector or an antibody raised against P1 is useful for  
CC treating diseases which can be ameliorated or cured by decreasing the  
CC level of the CL protein. The CL nucleic acids may be used to detect and  
CC quantitate expression of CL protein in biopsied tissues. The CL nucleic  
CC acids may also be used in methods for diagnosing diseases resulting from  
CC inherited defective CL sequences. The antibodies are also useful for  
CC detecting the CL protein in biological samples. This sequence encodes a  
CC human CXC family chemokine-like protein which is described in the method  
CC of the invention.  
XX Sequence 2575 BP; 517 A; 744 C; 629 G; 685 T; 0 other;  
SQ

Query Match 2.7%; Score 28; DB 21; Length 2575;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 669 ccatgttgaccagcgtgctcgaactc 696  
|||||  
DB 1340 ccatgttgaccagcgtgctcgaactc 1367

RESULT 45  
AAH18187  
ID AAH18187 standard; cDNA; 2797 BP.  
XX AAH18187;  
AC  
XX 26-JUN-2001 (first entry)  
DT  
XX Human cDNA sequence SEQ ID NO:18096.  
DE  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
KW  
XX Homo sapiens.  
OS  
XX EP1074617-A2.  
PN  
XX 07-FEB-2001.  
PD  
XX 28-JUL-2000; 2000EP-0116126.  
PF  
XX 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
PI  
XX WPI: 2001-318749/34.  
DR  
XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX Claim 8; SEQ ID 18096; 2537pp + CD ROM; English.  
PS The present invention describes primer sets for synthesising 5602  
XX full-length cDNAs defined in the specification. Where a primer set  
CC

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX Sequence 2797 BP; 621 A; 721 C; 655 G; 800 T; 0 other;  
SQ

Query Match 2.7%; Score 28; DB 22; Length 2797;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 669 ccatgttgaccagcgtgctcgaactc 696  
|||||  
DB 1529 ccatgttgaccagcgtgctcgaactc 1556

RESULT 46  
AAA71607  
ID AAA71607 standard; DNA; 2799 BP.  
XX AAA71607;  
AC  
XX 11-DEC-2000 (first entry)  
DT  
XX Human CXC family chemokine-like protein encoding DNA SEQ ID NO: 2.  
DE  
XX Human; CXC family; chemokine-like; disease; treatment; CL protein; ds.  
KW  
XX Homo sapiens.  
OS  
XX WO2000044784-A1.  
PN  
XX 03-AUG-2000.  
PD  
XX 21-JAN-2000; 2000WO-IL00043.  
PF  
XX 27-JAN-1999; 99IL-0128252.  
PR  
XX (COMP-) COMPUGEN LTD.  
PA  
XX Mintz L, Savitzky K;  
PI  
XX WPI: 2000-491238/43.  
DR  
XX P-FSDS: AAB10492.  
DR  
XX New human chemokine like (CL) nucleic acids and proteins, useful for  
PT diagnosing and treating diseases resulting from inherited defective CL  
PT sequences -  
XX Claim 1(i); Page 48-49; 55pp; English.  
PS This invention describes novel human chemokine like (CL) nucleic acids  
CC (N1), their complementary sequences (N2) and encoded amino acids (P1).  
CC An expression vector or P1 is useful for treating diseases which  
CC can be ameliorated or cured by raising the level of the CL protein. N2,  
CC the expression vector or an antibody raised against P1 is useful for  
CC treating diseases which can be ameliorated or cured by decreasing the  
CC level of the CL protein. The CL nucleic acids may be used to detect and  
CC quantitate expression of CL protein in biopsied tissues. The CL nucleic  
CC acids may also be used in methods for diagnosing diseases resulting from  
CC inherited defective CL sequences. The antibodies are also useful for  
CC detecting the CL protein in biological samples. This sequence encodes a  
CC human CXC family chemokine-like protein which is described in the method  
CC of the invention.

CC treating diseases which can be ameliorated or cured by decreasing the  
 CC level of the CL protein. The CL nucleic acids may be used to detect and  
 CC quantitate expression of CL protein in biopsied tissues. The CL nucleic  
 CC acids may also be used in methods for diagnosing diseases resulting from  
 CC inherited defective CL sequences. The antibodies are also useful for  
 CC detecting the CL protein in biological samples. This sequence encodes a  
 CC human CXC family chemokine-like protein which is described in the method  
 CC of the invention.

XX Sequence 2799 BP; 539 A; 825 C; 707 G; 728 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 2799;  
 Best Local Similarity 100.0%; Pred. No. 0.00046;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 669 ccatgttgaccaggctgtctcgaactc 696  
 Db 1564 ccatgttgaccaggctgtctcgaactc 1591  
 |||||

RESULT 47  
 AAHL7249  
 ID AAHL7249 standard; cDNA; 2845 BP.  
 AC AAHL7249;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:16639.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

FN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

PS Claim 8; SEQ ID 16639; 2537bp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 2845 BP; 839 A; 628 C; 661 G; 717 T; 0 other;

Query Match 2.7%; Score 28; DB 22; Length 2845;  
 Best Local Similarity 100.0%; Pred. No. 0.00046;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 663 gtttcgcatgttgaccaggctggtctc 690  
 Db 195 gtttcgcatgttgaccaggctggtctc 222  
 |||||

RESULT 48  
 AAH21279/c  
 ID AAH21279 standard; DNA; 3585 BP.

XX AC AAH21279;

XX 14-MAR-2001 (first entry)

XX Human low adenosine antisense oligonucleotide related sequence #2846.

KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
 KW human; airway disorder; bronchoconstriction; lung inflammation;  
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;  
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
 KW cancer; ss.

XX Homo sapiens.

XX WO200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08020.

XX 06-APR-1999; 99US-0127958.

XX (UYEC-) UNIV EAST CAROLINA.

XX (NYCE/) NYCE J W.

XX Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not  
 PT trigger adenosine receptors during metabolism, useful e.g. for treating  
 PT cancers and respiratory obstructions -

PS Disclosure; Page 1258-1259; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense  
 CC oligonucleotides and compositions (I) comprising them. In the antisense  
 CC oligonucleotides the A is replaced by a 'universal' or alternative base.  
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.  
 CC The antisense oligonucleotides and (I) can be used to down-regulate the

expression and or activity of target polypeptides associated with  
lung/respiratory disorders and malignancies, such as stimulating and  
activating peptide factors and transmitters, transcription factors,  
immunoglobulins and antibodies, antibody receptors, cytokines and  
chemokines, endogenously produced specific and non-specific enzymes,  
binding proteins, adhesion molecules and their receptors, cytokine and  
chemokine receptors, adenosine receptors, bradykinin receptors, central  
nervous system (CNS) and peripheral nervous and non-nervous system  
receptors, CNS and peripheral nervous and non-nervous system peptide  
transmitters, defensins, growth factors, vasoactive peptides and  
receptors, binding proteins and malignancy associated proteins. The  
antisense oligonucleotides may be used in this way to treat disorders  
including respiratory obstruction (especially pulmonary obstruction  
and/or bronchoconstriction) and/or lung inflammation, allergy(ies)  
and/or surfactant hypoproduction which are associated with a disease or  
condition selected from pulmonary vasoconstriction, inflammation,  
allergies, asthma, impeded respiration, respiratory distress syndrome  
(RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
pulmonary transplantation rejection, pulmonary infections, bronchitis,  
and/or cancer. AA18434 to AA21543 represent human polynucleotide  
fragments and antisense oligonucleotides used in the exemplification of  
the present invention.

Sequence 3585 BP: 864 A; 945 C; 1066 G; 710 T; 0 other:

Query Match 2.7%; Score 28; DB 21; Length 3585;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 catgttgaccaggctggtctcgaaactc 696  
|||||  
Db 1506 CCATGTTGACCAGGCTGGTCTCGAACTC 1479

RESULT 49  
AA335157/C  
ID AA335157 standard; DNA: 3585 BP.  
AC AA335157;  
28-JUL-2000 (first entry)  
Human adenosine receptor related polynucleotide 2nd SEQ ID NO:31.  
Human; adenosine receptor; low adenosine antisense oligonucleotide;  
phosphorothioate; impaired respiration; inflammation; allergy;  
allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;  
lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
Homo sapiens.  
OS  
XX  
XX  
PN WO200009525-A2.  
24-FEB-2000.  
XX  
XX  
PF 03-AUG-1999; 99WO-US17712.  
XX  
XX  
PR 03-AUG-1998; 98US-0095212.  
XX  
XX  
PA (UYEC-) UNIV EAST CAROLINA.  
XX  
XX  
PI Nyce JW;  
XX  
XX  
DR WPI; 2000-205971/18.  
XX  
XX  
PT New antisense oligonucleotides useful for treating e.g. pulmonary  
vasoconstriction, inflammation, allergies, asthma, hypertension,

bronchitis, emphysema, respiratory distress syndrome, ischemia or  
cancers -  
Disclosure: Page 1175-1176; 1343pp; English.  
The present invention describes a new composition comprising an  
antisense oligonucleotide (ON) with low adenosine (up to 15%), which  
targets nucleic acids involved in bronchoconstriction, allergies, and/or  
inflammation. The ON can have antiinflammatory, antiallergic,  
antiasthmatic, cytostatic and analgesic activities. The compositions are  
useful for the treatment of diseases associated with inflammation,  
impaired airways, including lung disease and diseases whose secondary  
effects afflict the lungs of a subject. They can be used for treating  
e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, cystic  
asthma, impeded respiration, respiratory distress syndrome, pain, cystic  
fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,  
carcinomas, and cancers which may metastasize to the lungs, including  
breast and prostate cancer. The reduction of the adenosine content of  
the ONs reduces side effects. The A-containing ONs break down with the  
release of deoxyadenosine which activates adenosine receptors causing  
bronchoconstriction and inflammation. AA33213 to AA33512 represent the  
nucleotide sequences given in the sequence listing from the present  
invention, which correspond to SEQ ID NO:1 to 2815, and then the last  
185 sequences are also called SEQ ID NO:1 to 185, but the sequences  
differ from the previously named sequences. SEQ ID NO:11 to 1680  
(AA33223 to AA33992) are specifically claimed ONs from the present  
invention. N.B. Sequences given in the disclosure of the present  
invention do not match up with their corresponding SEQ ID NO: sequences  
given in the sequence listing.

Sequence 3585 BP: 864 A; 945 C; 1066 G; 710 T; 0 other;

Query Match 2.7%; Score 28; DB 21; Length 3585;  
Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 catgttgaccaggctggtctcgaaactc 696  
|||||  
Db 1506 CCATGTTGACCAGGCTGGTCTCGAACTC 1479

RESULT 50  
AAH23153/C  
ID AAH23153 standard; DNA: 3586 BP.  
AC AAH23153;  
17-SEP-2001 (first entry)  
Nitric oxide synthase gene partial sequence.  
XX  
XX  
KW Hypertension; renal disease; non insulin dependent diabetes mellitus;  
polymorphism; nitric oxide synthase; breast; lung; prostate; cancer;  
KW nephrotropic; vasodilatory; renal injury; NIDDM; SNP; NOS; ds.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT promoter 1465..3472  
FT CDS /\*tag= a 3494..3586  
FT /\*tag= b  
/\*note= "partial coding sequence"  
XX  
PN WO200153537-A2.  
XX  
XX  
PD 26-JUL-2001.  
XX  
XX  
PF 24-JAN-2001; 2001WO-US02260.  
XX  
XX  
PR 24-JAN-2000; 2000US-0177775.

```

PR 25-JUL-2000; 2000US-0220662.
XX
PA (DZGE-) DZGENES LLC.
XX
PI Moskowitz DW;
XX
DR WPI; 2001-451916/48.
DR P-FSDB; AAB85401.
XX
XX
PT The diagnosis of disorders, i.e. hypertension and non insulin dependent
PT diabetes mellitus comprises the identification of single
PT polymorphism(s) in a nitric oxide synthase gene -
XX
PS Claim 1; Page 104-106; 119pp; English.
XX
CC The invention relates to diagnosing a genetic predisposition to disease,
CC condition or disorder i.e. hypertension, non insulin dependent diabetes
CC mellitus (NIDDM), end stage renal disease due to hypertension and/or
CC NIDDM mellitus that involves detecting the presence or absence of single
CC nucleotide polymorphisms (SNP) in the nitric oxide synthase gene (NOS).
CC The method is also useful for the treatment or prophylaxis of breast,
CC lung or prostate cancer. The method allows the identification of people
CC with a predisposition to disorders, before the disease progresses to an
CC end stage. The present sequence represents a partial sequence of the
CC human nitric oxide synthase gene.
XX
SQ Sequence 3586 BP; 864 A; 945 C; 1066 G; 710 T; 1 other;

Query Match          2.7%; Score 28; DB 22; Length 3586;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ccattgtgaccaggctggtctcgaactc 696
   |||
Db 1506 CCATGTTGACGAGGCTGCTCGAACTC 1479

Search completed: May 2, 2002, 19:45:11
Job time: 3700 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 13:50:18 ; Search time 2919.35 Seconds  
(without alignments)  
5882.667 Million cell updates/sec

Title: US-09-671-050-9  
Perfect score: 1041  
Sequence: 1 atggaaagtatgaaaaatt.....aggtacttcgctcaaaagt 1041

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_om.\*

20: em\_or.\*

21: em\_ov.\*

22: em\_pat.\*

23: em\_ph.\*

24: em\_pl.\*

25: em\_ro.\*

26: em\_sts.\*

27: em\_sy.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htgo\_hum.\*

31: em\_htgo\_inv.\*

32: em\_htgo\_rod.\*

33: em\_htg\_hum.\*

34: em\_htg\_inv.\*

35: em\_htg\_rod.\*

36: em\_htg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Match	Length	DB	ID	Description
1	1041	100.0	1041	6	AX107720	Sequence
2	1025	98.5	1068	6	AX107714	Sequence
3	839	80.6	945	6	AX107722	Sequence
4	823	79.1	972	6	AX107716	Sequence
5	757	72.7	1083	6	AX166534	Sequence
6	709	68.1	1819	6	AX107724	Sequence
7	537	51.6	911	6	AX056404	Sequence
8	454.4	43.7	561	6	AX107712	Sequence
9	454.4	43.7	594	6	AX107718	Sequence
10	453.2	43.5	2615	6	AX056405	Sequence
11	343.8	33.0	1363	9	HSSTHPKB	X66358 H.sapiens m
12	251.8	24.2	3153	4	AB029045	AB029045 Oryctolag
13	243.8	23.4	1518	10	AB029067	AB029067 Mus muscu
14	243.8	23.4	3988	10	AB029066	AB029066 Mus muscu
15	243.8	23.4	4022	10	AB029065	AB029065 Mus muscu
16	241.2	23.2	1993	9	HSU35146	U35146 Human p56 K
17	202.6	19.5	1773	9	AF130372	AF130372 Homo sapi
18	202.6	19.5	1773	6	AX166533	AX166533 Sequence
19	201	19.3	1974	3	AF326965	AF326965 Trypanoso
20	199	19.1	80362	9	AC079615	AC079615 Homo sapi
21	199	19.1	155028	33	AC037489	AC037489 Homo sapi
22	199	19.1	164281	2	AC092672	AC092672 Homo sapi
23	191	18.3	1732	10	AF112183	AF112183 Rattus no
24	191	18.3	1738	10	AF112184	AF112184 Rattus no
c 25	186.6	17.9	101340	2	AC018104	AC018104 Drosophil
c 26	186.6	17.9	190668	3	AC007419	AC007419 Drosophil
27	186.6	17.9	259973	3	AE003611	AE003611 Drosophil
28	177.2	17.0	3399	9	HSTXP3	Y15057 Homo sapien
29	175.2	16.8	1866	6	AX056407	AX056407 Sequence
30	166	15.9	960	3	TCU69958	U69958 Trypanosoma
31	163.8	15.7	903	8	AF026545	AF026545 Pneumocys
32	162	15.6	1476	6	AR087351	AR087351 Sequence
33	162	15.6	1476	9	HSCDK2MR	X61622 H.sapiens C
34	162	15.6	1635	6	A84437	A84437 Sequence 11
35	162	15.6	1635	6	A84499	A84459 Sequence 11
36	161.6	15.5	1297	9	HSCDK2	X62071 H.sapiens c
37	161	15.5	921	3	DDICRP	L00652 Dictyostel
38	160.4	15.4	2213	9	HUMCDC2A	M68520 Human cdc2-
39	160.4	15.4	2264	9	BC003065	BC003065 Homo sapi
40	155.6	14.9	2213	10	CGR223949	AJ223949 Cricetulu
41	154.6	14.9	885	3	AF076534	AF076534 Cryptospor
42	154.2	14.8	1050	9	HSCDC2	X05360 Human CDC2
43	154.2	14.8	1050	9	HSCDC2R	Y00272 Human cell
44	153	14.7	3321	3	LME293284	AJ293284 Leishmani
45	152.4	14.6	1170	5	AF159158	AF159158 Rana dybo

ALIGNMENTS

RESULT 1						
AX107720	AX107720	Sequence	9	1041 bp	DNA	PAT 30-APR-2001
LOCUS	Sequence	9	from Patent WO0123579.			
DEFINITION	AX107720					
ACCESSION	AX107720					
VERSION	AX107720.1	GI:13923201				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE	1 (bases 1 to 1041)					
JOURNAL	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and					
FEATURES	Sands,A.T.					
source	Human kinase proteins and polynucleotides encoding the same					
	Patent: WO 0123579-A 9 05-APR-2001;					
	Lexicon Genetics Incorporated (US)					
	Location/Qualifiers					
	1..1041					
	/organism="Homo sapiens"					

BASE COUNT 344 a 190 c 228 g 279 t  
ORIGIN

/db\_xref="taxon:9606"

Query Match 100.0%; Score 1041; DB 6; Length 1041;  
Best Local Similarity 100.0%; Pred. No. 1.2e-234;  
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 atgaaaaagatgaaaaattagctaaagctggagaaggtcttatggggtgtattcaaa 60
DB 1 ATGAAAAGATGAAAAATTAGCTAAGACTCGAGAAGGCTCTTATGGGTTGTATTCAAA 60
QY 61 tgcagaaacaaacctctggacaagtagtagctgttaaaaaatttggtaactcgaagat 120
DB 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAAATTTGTGGAAATCTGAAGAT 120
QY 121 gatcctgtgttaagaaaatagcactaaagaaatacgtatgttgaagcaattaaaacat 180
DB 121 GATCCTGTGTGTTAGAAAATAGCATAAGAGAATACGTATGTTGAAGCAATTAACAAT 180
QY 181 ccaaatcttgtgaacctcatcgaggtgttcaggagaaaaagaaatgcattagtttt 240
DB 181 CCAAAATCTGTGAACCTCATCGAGGTGTTCAGGAGAAAAAGGAAATGCATTTAGTTTT 240
QY 241 gaactctgatcatcacacttttaaatgagctggaagaaaccccaaatggagttcgtgat 300
DB 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTGTGTGAT 300
QY 301 ggagtgataaagcgtattatggaacacacttcaagctcttaattcttgcatacatat 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGCATATACAT 360
QY 361 aactgtattcacagagatataaaacctgaaaattcttaataaactaaagaagaataatc 420
DB 361 AACTGTATTACAGAGATATAAACCTTGAAAATATTCTTAATACTAAGCAAGGAATAATC 420
QY 421 aaagatttgacttcgggttgcacaaattcttgatccagagagatgctcacaccattat 480
DB 421 AAAGATTGTGACTTCGGGTTGCAAAATTTCTGATTCCAGGAGATGCCTACACCGATTAT 480
QY 481 gtgactcagagatggtaccagactcctgaaactcttgggagatactcagtagttct 540
DB 481 GTAGCTACGAGATGCTACCGAGCTCCTGGAACCTCTGTGTGGAGATACCTAGTATGGTTCT 540
QY 541 tcaatcagatataggctatggttggtttttggagagctcctgacagggccagccactg 600
DB 541 TCAGTCAATATATGGGCTATTGGTTGTTTGTGGAGAGCTCCTGACAGGCCAGCCACTG 600
QY 601 tggcctgaaaaatcagatgtggacaaactttatctgataatcagaacactagtagagacg 660
DB 601 TGGCCTGGAATAATCAGATGTGGACCAACTTTATCTGATAATCAGAACACTAGTAGAGACG 660
QY 661 ggggttcgcactgtgaccagctggtctgaaactcttgagacttgagctcaagtgatccactgccc 720
DB 661 GGGTTTCGCCATGTTGACAGGCTGGTCTCGAACTCTTGGAGTCAAGTGAATCCACCTGCC 720
QY 721 gtacaccttcaagtcgtggaattacagagaaaattaa'tcccaagaacatcaatcaatttt 780
DB 721 GTACCTCTCAAGTCTGGAAATTAACAGGAAAAATTAATCCCAAGACATCAATCAATCTTT 780
QY 781 aaaagtaacgggtttttccatggcatcagatataactgagccagaagaacatggaactctt 840
DB 781 AAAAGTAACGGGTTTTCCATGGCATCAGTATACCTGAGCCAGAAACATGGAACATCTT 840
QY 841 gaggaagaagtctcagatggttcatcctgtggctcgtgaacttcattgaaggggtgtctgaag 900
DB 841 GAGGAAAAGTCTCAGATGTTTCATCTCTGTGGCTGTGAACCTCATGAAGGGGTGTCGAAG 900
QY 901 atgaatccagatgacagattaaacctgttccaaactcctgaaagactcctcacttggattct 960
DB 901 ATGAATCCAGATGACAGATTAACTCTGTCCCAACTCTCTGGAGAGCTCTCTACTTTGATTCT 960
```

```
QY 961 ttccaagaggcccaaatataaagaaaacacgctaataatgaaggagaaaacagaagcgcacaa 1020
DB 961 TTCCAAGAGGCCCAAAATTAAGAAGAAAGCACGTAATGAAGGAAGAAACAGAAAGACGCCAA 1020
QY 1021 caggtaacttcgcgtcaaaaagt 1041
DB 1021 CAGGTACTTCGGCTCAAAAAGT 1041
```

```
RESULT 2
AX107714 LOCUS AX107714 1068 bp DNA PAT 30-APR-2001
DEFINITION Sequence 3 from Patent WO0123579.
ACCESSION AX107714
VERSION AX107714.1 GI:13923198
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 1068)
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and
Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 3 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Source 1..1068
/organism="Homo sapiens"
BASE COUNT 344 a 197 c 240 g 287 t
ORIGIN
```

Query Match 98.5%; Score 1025; DB 6; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 7.1e-231;  
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 atgaaaagatgaaaaattagctaaagctggagaaggtcttatggggtgtattcaaa 60
DB 1 ATGAAAAGATGAAAAATTAGCTAAGACTGGAGAAGGCTTATGGGTTGTATTTCAAA 60
QY 61 tgcagaaacaaacctctggacaagtagtagctgttaaaaaatttggaaatcgaagat 120
DB 61 TGCAGAAACAAACCTCTGGACAAGTAGTAGCTGTAAATAAATTTGTGGAATCTGAAGAT 120
QY 121 gatcctgtgttaagaaaaatagcactaaagaaatacgtatgttgaagcaattaaaacat 180
DB 121 GATCCTGTGTTTGAATAATAGCACCTAAGAGAAATAGCTATGTTGAAGCAATTAACAAT 180
QY 181 ccaaatctgtgaacctcatcgaggtgttcaggagaaaaagaaaatgcatttagtttt 240
DB 181 CCAAAATCTGTCAACCTCATCGAGGTGTTTCAGGAGAAAAAGGAAATGCATTTAGTTTT 240
QY 241 gaactctgatcatcacacttttaaatgagctggaagaaaccccaaatggagttcgtgat 300
DB 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTGTCTGAT 300
QY 301 ggaagtgtcaaaaacgctattatgccaacacttcaagctcttaattctgtcatatacat 360
DB 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAAGCTCTTAATTTCTGTCATATACAT 360
QY 361 aactgtattcacagagatataaaacctgaaaattcttaataaactaaagaagaataatc 420
DB 361 AACTGTATTCACAGAGATATAAACCTGAAATATTCTTAATACTAAGCAAGGAATAATC 420
QY 421 aaagatttgacttcgggttgcacaaattctgattccaggagatgcctacaccgattat 480
DB 421 AAAGATTGTGACTTCGGGTTGCAAAATTTCTGATTCCAGGAGATGCCTACACCGATTAT 480
QY 481 gtgactcagagatggtaccagactcctgaaactcttgggagatactcagtagttct 540
DB 481 GTAGCTACGAGATGTTACCGAGCTCCTGAACCTCTCTGGAGAGACTCTCAGTATGGTTCT 540
```



KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 972)  
AUTHORS Donoho, G., Turner, C.A., Nehls, M., Friedrich, G., Zambrowicz, B. and Sands, A.T.  
TITLE Human kinase proteins and polynucleotides encoding the same  
JOURNAL Patent: WO 0123579-A 5 05-APR-2001;  
FEATURES Lexicon Genetics Incorporated (US)  
Location/Qualifiers  
source I. .972  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 323 a 172 c 214 g 263 t  
ORIGIN  
  
Query Match 79.1%; Score 823; DB 6; Length 972;  
Best Local Similarity 90.6%; Pred. No. 1.8e-183;  
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;  
  
Qy 1 atggaaaagtataaaaaattagctaaagactggaggaagggtcttatgggtttattcaaa 60  
Db 1 ATGGAAGAGTATGAAAAATTAGCTAAGACTGGAGAGGGCTTATGGGGTTGATTCAAA 60  
Qy 61 tgcagaaacaaacctctggacaagtagtagctgtgttaaaaaatttggaaatctgaagat 120  
Db 61 TCGAAGAACAAACCTCTGGACAAGTAGTGTGTAAAAAATTTGTGAATCTGAAGAT 120  
Qy 121 gatcctgtgttaagaaaatagcactaaagaaatacgtatgttggaagcaataaaacat 180  
Db 121 GATCCTGTGTGTTAAGAAAATAGCATTAGAGAAATACGTATGTTGAAGCAATTAACAT 180  
Qy 181 ccaaatctgtgaacctcatcgaggtgttcagagaaagaaaggaagaaatgattgatttt 240  
Db 181 CCAAAATCTGTGAACCTCATCGAGGTGTTACAGAGAAAGGAAATGCAATTTAGTTT 240  
Qy 241 gaatactgtatcatactacttttaaatagctgtgaaagaaaccccaaatggagttgctgat 300  
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300  
Qy 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat 360  
Db 301 GGAGTGATCAAAAGCGTATTATGCAACACACTTCAAGCTCTTAATTTCTGTCATATACAT 360  
Qy 361 aactgtattcacagagatataaaacctgaaataattcttaataactaagcaaggaataatc 420  
Db 361 AACTGTATTACAGAGATATATAAACCTGAAATATTTCTAATACTAAGCAAGGAATAATC 420  
Qy 421 aagattgtgacttcgggtttgcacaaattctgattccagagagatgcctacacggattat 480  
Db 421 AAGATTGTGTGACTTCGGGTTTGACAAATTTCTGATTCAGAGAGATGCCTACACGATTAT 480  
Qy 481 gtactcaagagatgtaccagactctgaactctctgtgggagataactcaatggtttct 540  
Db 481 GTAGCTACGAGATGGTACCAGCTCTGAACTTCTGTGGAGATACTCAGTATGGTCT 540  
Qy 541 tcagtcgatataatgggctatgggtgtgttttgcagagactctctgacaggccagagcaactg 600  
Db 541 TCAGTCGATATATGGCTATTTGGTTGTGTTTTGCAGAGCTCCTGCACAGGCCAGCCACTG 600  
Qy 601 tggccttgaaaaatcagatgtggaccaactcttctgataatcagaacactagtagagacg 660  
Db 601 TGGCCTTGAAAAATCAGATGTGGACCACTTTATCTGATTAATCAGACACT----- 650  
Qy 661 gggtttcgcaatgttgaccaggtggtctcgaaactcttgaaagtgaaagtgatccacctgac 720  
Db 661 -----  
Qy 721 gtgacctcaaaagtgtggaattacaggaaaataatcccaagacatcaatcatttt 780  
Db 721 -----

Db 651 -----AGGAAAAATTATCCCAAGACATCAATCAATCTTT 684  
Qy 781 aaagtaaacgggtttttccatggcatcagttacacctggagccagagacatggaaactctt 840  
Db 685 AAAAGTAAACGGGTTTTTCATGGCATCAGTATACCTGAGCCAGAGACATGGAAACTCTT 744  
Qy 841 gaggaaaagtctcagatgttcacctctgtgctctgaaacttcataagagggtgtctgaag 900  
Db 745 GAGCAAAAGTTCTCAGATGTTTCATCTGTGGCTCTGAACCTTCATGAAGGGGTGTCTGAAG 804  
Qy 901 atgaatccagatgacagattaacctgttcccaactcctctggagagctcctacttattct 960  
Db 805 ATGAATCCAGATGACAGATTAACTGTGCCAACTCTGGAGAGCTCTACTATTGATCT 864  
Qy 961 ttccaagagcccaaatataaagaaagcacgtaatgaagaaagaaacagaaagacgcaa 1020  
Db 865 TTTCAAGAGGCCAAATTAAGAAAGAACACGTAATGAAGGAGAAACAGACAGCGCAA 924  
Qy 1021 caggt 1025  
Db 925 CAGGT 929  
  
RESULT 5  
LOCUS AX166534 1083 bp DNA PAT 22-JUN-2001  
DEFINITION Sequence 25 from Patent WO0138503.  
ACCESSION AX166534  
VERSION AX166534.1 GI:14546879  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1083)  
AUTHORS Plowman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R., Flanagan, P. and Clary, D.S.  
TITLE Novel human protein kinases and protein kinase-like enzymes  
JOURNAL Patent: WO 0138503-A 25 31-MAY-2001;  
FEATURES  
source Location/Qualifiers  
1. .1083  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 366 a 199 c 225 g 293 t  
ORIGIN  
  
Query Match 72.7%; Score 757; DB 6; Length 1083;  
Best Local Similarity 86.1%; Pred. No. 5.6e-168;  
Matches 927; Conservative 0; Mismatches 0; Indels 150; Gaps 2;  
  
Qy 1 atggaaaagtataaaaaattagctaaagactggaggaagggtcttatgggtttattcaaa 60  
Db 1 ATGGAAGAGTATGAAAAATTAGCTAAGACTGGAGAGGGCTTATGGGGTTGATTCAAA 60  
Qy 61 tgcagaaacaaacctctgacaaagtagtagctgtttaaaaaatttggaaatctgaagat 120  
Db 61 TCGCAAAACAAACCTCTCGACACAGTAGTAGCTGTTAAAAAATTTGTGGAATCTGAAGAT 120  
Qy 121 gatcctgtgttaagaaaaatagcactaagagaaatacgtatgttgaagcaattaaaaaat 180  
Db 121 GATCCTGTGTGTTAAGAAAAATAGCACTAAGAGAAATACGTATGTTGAAGCAATTAACAT 180  
Qy 181 ccaaatctcttgaacctcatcgaggtgttcagagaaagaaagaaatgattagttttt 240  
Db 181 CCAAAATCTGTGAACCTCATCGAGGTGTTCAGAGAAAGGAAATGCAATTTAGTTT 240  
Qy 241 gaatactgtatcatactacttttaaatagctgtgaaagaaaccccaaatggagttgctgat 300  
Db 241 GAATACTGTGATCATACACTTTTAAATGAGCTGGAAGAAACCCAAATGGAGTTGCTGAT 300  
Qy 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat 360



Db	301	CGAGTGATCAAAAGCGTATTATGCGAACAACCTTCAAGCTCTTAATTCTGTCTATACAT	360
Qy	361	aactgtattacagagataataaacctgaaatatattctaataactaacaaggaataatc	420
Db	361	AACGTGATTACAGAGATATAAAACCTGAAATATTCATAAATAACAGAGGAATATC	420
Qy	421	aagatttgacttcgggttgacaaattctga	454
Db	421	AAAGATTGTGACTTCGGGTTTGACAAATTCAGTAATTCACATGTTGGGAAATTGAC	480
Qy	455	-----ttccagagagatgacctacacccattatgact	486
Db	481	TTACTGATCTGATGTCATTGTCAGTTTCCAGAGATGCCTACACCGATTATGAGCT	540
Qy	487	acgagatggttacagagctcctgaaactctgtggtgagatactcagtatggtcttcagtc	546
Db	541	ACGAGATGTTACCGAGCTCCTGAACCTTCTGTGGGAGATACCTCAGTATGTTCTTCAGTC	600
Qy	547	gatataaggctattggtgtgttttgcagagctcctgcagagccagccactgtggcct	606
Db	601	GATATATGGGCTATTGGTTGTGTTTTTGCAGAGCTCCTGCAGGCCAGCCACTGTGGCCT	660
Qy	607	ggaataatcagatggaacaacttatactgataatacagaacactagtagagacggggttt	666
Db	661	CGAAATCAGATGGACCAACTTTATCTGATAATCAGAACACT	704
Qy	667	cgcattgtgacagcgtggtctcgaaactctgaogtcaagtgatccacctgcgctagacc	726
Db	705	-----	704
Qy	727	tctcaagtgctggaattacaggaataataatcccaagacatcaatcaatctttaaagt	786
Db	705	-----AGGAAATTAATCCCAAGACATCAATCAATCTTTAAGT	744
Qy	787	aacgggttttccatcagtcagtcacccagagacatggaacactcttgaggaa	846
Db	745	AACGGGTTTTTCCATGTCATCAGTATACCTGAGCCAGAGACATGGAACCTCTGAGGAA	804
Qy	847	aagttctcagatgctcctgctcctgaacttcataaggggtgctcgaagatgaat	906
Db	805	AGTTCTCAGATGTTATCTGTGGCTCTGAATTCATGAAGGGGTCTCTGAAGATGAAT	864
Qy	907	ccagatgacagattaacctgttcccaactcctggagcctcctacttgattcttttcaa	966
Db	865	CCAGATGACAGATTAACTGTTCCCAACTCCTGGAGAGCTCTACTTTGATTCTTTCAA	924
Qy	967	gagggccaaataaadaaagcagcgtaatgaaggaagaaacagagacgccacag	1023
Db	925	GAGGCCAAATTAAGAAAGACGTAATGAGGAGAAACAGAGAGCGCCACAG	981
RESULT 6			
AX107724			
LOCUS	AX107724	1819 bp	DNA
DEFINITION	Sequence 13 from Patent WO0123579.		
ACCESSION	AX107724		
VERSION	AX107724.1	GI:13923203	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1819)		
AUTHORS	Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.		
TITLE	Human kinase proteins and polynucleotides encoding the same		
JOURNAL	Patent: WO 0123579-A 13 05-APR-2001;		
FEATURES	Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers		
	1..1819		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		

BASE COUNT		539 a	368 c	406 g	505 t	1 others
ORIGIN						
Query Match		68.1%; Score 709; DB 6; Length 1819;				
Best Local Similarity		77.0%; Pred. No. 1.1e-156;				
Matches 1025; Conservative		0; Mismatches 0; Indels 306; Gaps 1;				
Qy	1	atggaaaagatgaaaaattagctaagactgagagaagggtcttatggggttgtattcaaaa	60			
Db	138	ATGGAAAAGATGAAAAATTAGCTAAGACTGGAGAAGGCTCTTATGGGGTTGTATTCAA	197			
Qy	61	tgcagaaaaacaaacctctggacaagtagctgttlaaaaaatttgggaactcgaagat	120			
Db	198	TGCAGAAAACAAACCTCTGGACAAGTAGTGTCTTAAAAAATTTGTGGAATCTGAAGAT	257			
Qy	121	gactcgtgtttaagaaaaatagcactaagagaaatcacgtatgttgagcaaatataaacat	180			
Db	258	GATCCTGTGTTTAAAGAAAATAGCACTAAGAGAAAATACGTATGTTGAAGCAATTAACACAT	317			
Qy	181	ccaaatcttgtgaacctcatcgagggtgttcagagagaaaaaagaaatgcattttagtttt	240			
Db	318	CCAAATCTTGTGAACCTCATCGAGGTGTTTCAGGAGAAAAGAAAATGCATTAGTTT	377			
Qy	241	gaatactgtgatacacacttttaaatgagctggaagaaaccccaaatggagtctctgat	300			
Db	378	GAATACTGTGATCCTACACTTTTAAATGAGCTGGAAGAAACCCCAATGGAGTTGCTGAT	437			
Qy	301	ggagtgataaaagcgattatggcaaacacttcaagctcttaattctgtgcatacat	360			
Db	438	GGAGTGATCAAAAAGCGTATTATGGCAACACTTCAAGCTCTTTAAATTTCTGTCTATACAT	497			
Qy	361	aactgtattcacagagataataaacctgaaaatttctaataactaagaaggaataatc	420			
Db	498	AACTGTATTACAGAGATATAAACCCTGAAAATATTCTAATAACTAAGCAAGGAATAATC	557			
Qy	421	aagatttgtgaacttcgggttgcaaaaattctg	453			
Db	558	AAGATTTGTGACTTCGGGTTTGCACAAATCTGAGTTGGAGCTTCATCTTTCTGTGTGCC	617			
Qy	454	-----	453			
Db	618	TCCTTGATTGGCTTAATAGTTGACCTTCTGAATTTCTTTTCTGCCAATTCAGAGATTTT	677			
Qy	454	-----	453			
Db	678	CTCCTGGCTTGGATCCATTGCTGACACAGTGTTCACCATGGGCGCCAGGCTCATCTCGA	737			
Qy	454	-----	453			
Db	738	ACTTCTGGCCTCAAGTGATCTTCCACCTCGGCTCCCAAAAGTGTGATTGCAAGTGTG	797			
Qy	454	-----	453			
Db	798	AGCCACCGTGCCCGACGAGATTTTTCAAACAATAACTACTGAGAGCTCACAAGATTTT	857			
Qy	454	-----attccaggagatgcctacacc	474			
Db	858	TTAGTGGGACACAAATTTTCGAACAANAATCTTGAAGACGCAATTCAGGAGATGCCATCAC	917			
Qy	475	gattatgactacagatggttacccagctcctgaactcttctgtgggagatactcagtat	534			
Db	918	GATTATGATGACTACGAGATGGTACCGAGCTCCTGAACCTTCTTGTGGGAGATACTACGAT	977			
Qy	535	ggttcttcagtcgatatatgggtctattggtgtgttttttcagagctcctcagagccag	594			
Db	978	GGTTCTTCAGTCGATATATGGGCTATTGGTTGTGTTTTCAGAGAGCTCTCTGACAGGCCAG	1037			
Qy	595	ccactgtggcctggaaaaatcagatgtggaccaactttatcigataatcagaaactagta	654			
Db	1038	CCACTGTGCGCTGGAAAATCAGATGTGACCAACTTTATCTGATATAACAGACACTAGTA	1097			
Qy	655	gaagaggggtttccgcattgttgaccaggtggtctcgaactctctgcagctcaagtgatcca	714			

```
|||||
Db 1098 GAGACGGGTTTCGCATGTTGACGAGCTGCTCGAACTCTTGACGTCAAGTGATCCA 1157
QY 715 cctgcgtagcctcctcaagtctggaattacaggaataattatcccaagacatcaatca 774
Db 1158 CTFGCGCTAGCCTCTCAAAGTCTGGAATTACAGAAATTAATCCCAAGACATCAICA 1217
QY 775 atctttaaagaataacacgggttttccatggcatcagtatcctcgagccagagacatgaa 834
Db 1218 ATCTTTAAAGTAACGGGTTTTCCTGCGCATCAGTATACCTTGAGCCAGAGACATGGAA 1277
QY 835 actcttgaggaataagttctcagatttccctgtggtctggaacttcacatgaagggtgt 894
Db 1278 ACTCTTGAGGAAGTCTCAGATGTTTCATCCTGTGGCTCTGAACCTTCATGAAGGGGTG 1337
QY 895 ctgaagatgaatccagatgacagattaacctgttcccaactcctcgagagcctcacttt 954
Db 1338 CTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCCTGGAGAGCTCCTACTTT 1397
QY 955 gattctttcaagagcccaaatataaagaaaaagcagcgttaatgaaggaacagagaaga 1014
Db 1398 GATCTCTTTCAAGAGGCCCAATTAAGAGAAAGACGCTAATGAAGGAAGAACAGAGA 1457
QY 1015 cgccaacaggt 1025
Db 1458 CGCCAACAGGT 1468

RESULT 7
AX056404 911 bp DNA PAT 13-JAN-2001
DEFINITION Sequence 48 from Patent WO0073469.
ACCESSION AX056404
VERSION AX056404.1 GI:12229111
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 911)
AUTHORS Plozman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 48 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
source
1..911
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 293 a 182 c 178 g 258 t
ORIGIN

Query Match 51.6%; Score 537; DB 6; Length 911;
Best Local Similarity 87.0%; Pred. No. 2.4e-116;
Matches 643; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

QY 285 aatggaattgctgtagtgatgatacaaaagcgtattatggcaaacacttcagctctaa 344
Db 1 AATGGAGTTGCTGATGGATGATCAAAAGCGTATTATGGCAAACTTCAAGCTCTTAA 60
QY 345 ttctgtcatatcataactgtattcacagagatataaaacctgaaaatatcttaataac 404
Db 61 TTTCTGTATATACATAAATGATTTACAGAGATATATAAACCTGAAATATCTATATAC 120
QY 405 taagcaagaataatcaagatttgacttcgggtttgacaaattctgattccaggaga 464
Db 121 TAGCAAGGATATATCAAGATTGTGACTTCGGGTTCGCACAAATTTGATTCAGGAGA 180
QY 465 tgctacaccgattatgtagctacagatggtaaccagctcctcgaaacttctgtggaga 524
Db 181 TGCCATACACCGATTATGTAGCTACGAGATGTCACGAGCTCTCGAACTCTTGTGGGAGA 240
QY 525 tactcagtaggttcttcagtcgatatatggctattggtgtgtttttgcagagctcct 584
```

```
|||||
Db 241 TACTCAGTATGGTTCTTCAGTCGATATATGGGTATGTTGTTTTCGAGAGTCTCT 300
QY 585 gacagggccagcactctgctcgtggaataatcagatgtggaccaactttatctgataatcag 644
Db 301 GACAGGCCAGCCACTCTGGCCCTGGAAATCAGATGTGGACCAACTTTATCTGATAATCAG 360
QY 645 aacactagtagagacgagggtttcgccatgttgaccagcgtgctcgaacctcttgacgtc 704
Db 361 AACACT----- 366
QY 705 aagtgatccacctgcgtagcctcctcaagtgctggaattacaggaataataatcccaag 764
Db 367 -----AGGAAAAATTAATCCCAAG 384
QY 765 acatcaatcaatctttaaagtaacgggtttttccatggcatcagtatccctgagccaga 824
Db 385 ACATCAATCAATCTTTAAAGTAACGGGTTTTCATGCGCATCACTATACCTTGAGCCAGA 444
QY 825 agacatggaactcttgaggaataatctcagatgttcatcctgtggtctgaaacttcac 884
Db 445 AGACATGGAAACTCTTTGAGGAAAGTTCTCAGATGTTTCATCCTGTGGCTCTGAACCTCAT 504
QY 885 gaagggtgtctgaagatgaatccagatgacagatgaacacctgttcccaactcctgagag 944
Db 505 GAAGGGGTGTCTGAAGATGAATCCAGATGACAGATTAACCTGTTCCCAACTCTCTGGAGAG 564
QY 945 ctctacttgattctttcaagagcccaaatataaagaaagcagcgttaatgaaggaag 1004
Db 565 CTCTACTTGTGATCTTTCAAGAGGCCCAATTAAGAGAAAGACGCTAATGAAGGAAG 624
QY 1005 aaacagaagacgccaacag 1023
Db 625 AAACAGAAGACGCCCAACAG 643

RESULT 8
AX107712 561 bp DNA PAT 30-APR-2001
LOCUS Sequence 1 from Patent WO0123579.
DEFINITION AX107712
ACCESSION AX107712.1 GI:13923197
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 561)
AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.
TITLE Human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0123579-A 1 05-APR-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
1..561
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 189 a 87 c 112 g 173 t
ORIGIN

Query Match 43.7%; Score 454.4; DB 6; Length 561;
Best Local Similarity 99.8%; Pred. No. 5.7e-97;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggaaaagtataaaaattagcttaagactggagaagggtcttatggggttgattacaaa 60
Db 1 ATGGAAAAGTATGAAAATTTAGCTAAGACTGGAGAAGGGTCTTATGGGGTGTGTTATCAA 60
QY 61 tgcagaacaaaaacctctggacaagtagtgctgttaaaaaatttgggaactcgaagat 120
Db 61 TCGAACAACAAAACCTCTGGACAAGTAGTAGCTTTAAAAAATTTGTGGAACTCGAAGT 120
```

QY 121 gatcctgttgaataaataagcactaagagaaatacgtatgttgaaagcaataaacaat 180  
 Db 121 GATCCTGTTCTTAAGAAATAGCAGCTAAGAGAAATACGTATGTGGAAGCAATAAACAT 180  
 QY 181 ccaaatctgtgaacctcatcgaggtgttcagagagaaaaaggaatgcatttagtttt 240  
 Db 181 CCAATCTTGTGAACCTCATCGAGTGTTCAGAGAGAAAAGGAATGCTATTAGTTTTT 240  
 QY 241 gaalactgtgatacacacttttaaatgagctggaaagaaaccccaaatggagttcgtgat 300  
 Db 241 GAATACTGTGATCATACACTTTTAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT 300  
 QY 301 ggaagtatacaagagctattatgcaaacactcoagctcttaattcttgcatacatat 360  
 Db 301 GGAGTGTATCAAAAGCGTATTATGGCAACACTTCAGCTCTTAATTTCTGTCATATACAT 360  
 QY 361 aactgtattcacagagatatataaacctgaaatatattcttaataactaagaagaataatc 420  
 Db 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTTAATACTAAGCAAGGAATAATC 420  
 QY 421 aagatttgtgacttcgggtttgcacaaattctgatt 456  
 Db 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGAGT 456  
 RESULT 9  
 AX107718 594 bp DNA PAT 30-APR-2001  
 LOCUS AX107718 Sequence 7 from Patent WO0123579.  
 DEFINITION AX107718  
 ACCESSION AX107718  
 VERSION AX107718.1 GI:13923200  
 KEYWORDS human.  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 AUTHORS Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and  
 Sands,A.T.  
 TITLE Human kinase proteins and polynucleotides encoding the same  
 JOURNAL Patent: WO 0123579-A 7 05-APR-2001;  
 Lexicon Genetics Incorporated (US)  
 FEATURES  
 source Location/Qualifiers  
 1. .594  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 203 a 94 c 115 g 182 t  
 ORIGIN  
 Query Match 43.7%; Score 454.4; DB 6; Length 594;  
 Best Local Similarity 99.8%; Pred. No. 5.7e-97;  
 Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 atgaaaaagtatgaaaaattagctaagactggagaaggggtcttatggggtgtattcaaa 60  
 Db 1 ATGAAAAGTATGAAAATTTACCTAAGACTGGAGAGGGTCTTATGGGTTGTATTCAAA 60  
 QY 61 tcagaaacaaacctctggacaagttagctgttaaaaaatttggtaaatctgaagat 120  
 Db 61 TCAGAAAACAAACCTCTGGACAAGTAGTACTGTTAAAAAATTTGTGGAATCTGGAAGAT 120  
 QY 121 gatcctgttgaagaaatagcactaagagaaatacgtatgttgaaagcaataaacaat 180  
 Db 121 GATCTGTGTTGAAGAAATAGCAGCTAAGAGAAATACGTATGTGGAAGCAATTAACAT 180  
 QY 181 ccaaatcttgtgaacctcatcgaggtgttcagagagaaaaaggaatgcatttagtttt 240  
 Db 181 CCAAAATCTTGTGAACCTCATCGAGGTGTTCAGAGAGAAAAGGAATGCATTAGTTTTT 240  
 QY 241 gaatactgtgatacacacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300  
 Db 241 GAATACTGTGATCATACACTTTTAATGAGCTGGAAAGAAACCCCAATGGAGTTGCTGAT 300

QY 301 ggagtgatcaaaagcggtattatggcaaacactcoagctcttaattctgtcatatacat 360  
 Db 301 GGAGTGATCAAAAGCGTATTATGGCAACACTTCAGCTCTTAATTTCTGTCTATATACAT 360  
 QY 361 aactgtattcacagagatatataaacctgaaatatattcttaataactaagaagaataatc 420  
 Db 361 AACTGTATTACAGAGATATAAAACCTGAAATATTTCTTAATACTAAGCAAGGAATAATC 420  
 QY 421 aagatttgtgacttcgggtttgcacaaattctgatt 456  
 Db 421 AAGATTGTGACTTCGGGTTTGCACAAATTTCTGAGT 456  
 RESULT 10  
 AX056405 2615 bp DNA PAT 13-JAN-2001  
 LOCUS AX056405 Sequence 49 from Patent WO0073469.  
 DEFINITION AX056405  
 ACCESSION AX056405  
 VERSION AX056405.1 GI:12229112  
 KEYWORDS Murinae gen. sp.  
 SOURCE Murinae gen. sp.  
 ORGANISM Murinae gen. sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.  
 REFERENCE 1 (bases 1 to 2615)  
 AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.  
 TITLE Protein kinases  
 JOURNAL Patent: WO 0073469-A 49 07-DEC-2000;  
 Sugen, Inc. (US)  
 FEATURES  
 source Location/Qualifiers  
 1. .2615  
 /organism="Murinae gen. sp."  
 /db\_xref="taxon:39108"  
 BASE COUNT 742 a 580 c 585 g 708 t  
 ORIGIN  
 Query Match 43.5%; Score 453.2; DB 6; Length 2615;  
 Best Local Similarity 73.1%; Pred. No. 1.2e-96;  
 Matches 658; Conservative 0; Mismatches 143; Indels 99; Gaps 2;  
 QY 135 gaaaatgactaagaagaaatcgtatgttgaagcaattaaacacatccaaatcttgtgaa 194  
 Db 1 GAAAATAGCTCTCGGGAATCCGTATGCTGAAG---TTGAACAACCAACCTCTGTGA 57  
 QY 195 cctcatcgaggtgttcaggagaaaaagaaatgcatttagtttttgaatactgtgata 254  
 Db 58 CCTCATCGAGGTGTTGAGAAGAGAGAAAGATGCATCTAGTTTTTGTAGTACTGTGATCA 117  
 QY 255 tacatttttaaatgagctggaagaaaccccaaatgaggtgctgagtgatgatacaaaag 314  
 Db 118 CACATGTTAAACGAGCTGGAGAGAAACCCAAACGGAGTTTCTGATGAGTGATTAAAG 177  
 QY 315 cgtattatggcaaacattcaagctcttaattctgtcatatacacataactgtattcacag 374  
 Db 178 TGTGTTATGCAAAACCCCTTAAGCCCTTAACCTCTGTCTCAACAGACAAATTTGTTATCG 237  
 QY 375 agatataaaacctgaaatatattcttaataactaagaagaataatcaagatttgtgact 434  
 Db 238 GGAATGTAACCTGAAACATCCTTAATTAACCAACGAAGGATGATAAAGATTTGTGACTT 297  
 QY 435 cgggtttgcacaaattctgattccaggagatgctcacaccagattatgtagctacagatg 494  
 Db 298 TGGATTTCACAGAAATTTAATTTCCAGAGAGCCCTACACAGACTATGTTGCCACAGGGTG 357  
 QY 495 gtaccgagctcctgaacctctgttggagatactcagtagtcttcttcagtcgatatatg 554  
 Db 358 GTACCGAGCCCCCAACTTCTCGTGGGACACCAAGTAGCGGTCTCTGTAGACGTGIG 417  
 QY 555 ggctatgtgttgttttttgcagagagcctcctgacaggccagcgaactgtggcctggaaaaac 614  
 Db 418 GGCCGTGCGCTGTGTGTGTTCAGAGCTCTTCAGCGGTCAGCCACTCTGCGCCGGGAAAAATC 477









.....



12	251.8	24.2	3080	21	AA29745	Rabbit KTIAMRE kin
13	202.6	19.5	1513	21	AA251208	Human lost in leuk
14	202.6	19.5	1773	22	AA506724	Polynucleotide seq
15	191	18.3	1667	21	AA251207	Rat lost in leukae
16	191	18.3	1738	21	AA251206	Rat lost in leukae
17	175.2	16.8	1866	22	AAF44672	Novel protein kina
18	163.8	15.7	903	20	AA07476	Pneumocystis carin
19	162	15.6	1476	21	AA237835	Cyclin-dependent k
20	162	15.6	1635	19	AAV71073	CDK2-green fluore
21	162	15.6	1635	19	AAV71074	Green fluorescent
22	160	15.4	1322	21	AAI5889	Human prostate can
23	134.8	12.9	1825	22	AAI58432	Human polynucleoti
24	133.2	12.8	1089	15	AAQ63490	Cyclin dependent p
25	132	12.3	1002	18	AA64448	CDK1 protein. Can
26	128	12.3	1002	18	AAZ08264	Candida gene encod
27	127	12.2	6015	21	AAZ75931	Human ORF ORF1486
28	125.6	12.1	1788	22	AAI60218	Human polynucleoti
29	125.4	12.0	1310	21	AAZ35344	Arabidopsis thalia
30	124.6	12.0	1899	22	AAZ44678	Novel protein kina
31	124.6	12.0	1293	21	AAZ36413	Arabidopsis thalia
32	123.8	11.9	1307	21	AAZ51531	Arabidopsis thalia
33	123.4	11.9	1326	21	AAZ43758	ze a may DNA fragm
34	116	11.1	1070	18	AAZ64450	MOC1 protein. Can
35	116	11.1	1070	18	AAZ08266	Candida gene encod
36	115.4	11.1	1353	21	AAZ46587	ze a may DNA fragm
37	110.8	10.6	996	22	AAH46905	CDNA encoding huma
38	107.2	10.3	936	22	AAF58252	Oligonucleotide D1
39	107.2	10.3	936	22	AAF58254	Oligonucleotide D1
40	107.2	10.3	936	22	AAF58257	Oligonucleotide D2
41	107.2	10.3	936	22	AAF58259	Oligonucleotide D2
42	107.2	10.3	936	22	AAF58262	Oligonucleotide D2
43	107.2	10.3	938	22	AAF58255	Oligonucleotide D1
44	106.2	10.2	936	22	AAF58252	Oligonucleotide D1
45	106.2	10.2	936	22	AAF58254	Oligonucleotide D1

ALIGNMENTS

RESULT	1
AD03816	
ID	AD03816 standard; cdna; 1041 BP.
AC	AD03816;
XX	AD03816;
XX	19-JUN-2001 (first entry)
XX	Human kinase cdna #5.
DE	Human; kinase: gene therapy; bioreactor; mental disorder;
KW	biological disorder; ss.
XX	Homo sapiens.
XX	Key
XX	Location/Qualifiers
FT	1..1041
FT	/tag= a
FT	/product= "Human kinase #5"
FT	/note= "The coding region does not include stop codon"
FT	/partial
XX	
PN	W0200123579-AL.
XX	
PD	05-APR-2001.
XX	
PF	27-SEP-2000; 2000WO-US26621.
XX	
PR	28-SEP-1999; 99US-0156511.
XX	
PA	(LEXI-) LEXICON GENETICS INC.
XX	Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
PI	
XX	

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 15:42:23 ; Search time 234.75 Seconds  
(without alignments)  
3801.813 Million cell updates/sec

Title: US-09-671-050-9

Sequence: 1 atggaagatgataaaatt.....aggtaacttcgctcaaaagt 1041

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :	N_Geneseq_1101.*
1:	/SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	1041	22	AD03816 Human kinase cdna
2	1025	98.5	1068	22	AD03813 Human kinase cdna
3	839	80.6	945	22	AD03817 Human kinase cdna
4	823	79.1	972	22	AD03814 Polynucleotide seq
5	757	72.7	1083	22	AA506725 Human kinase cdna
6	709	68.1	1819	22	AA03818 Novel protein kina
7	537	51.6	911	22	AAF44669 Human kinase cdna
8	454.4	43.7	561	22	AA03812 Human kinase cdna
9	454.4	43.7	594	22	AA03815 Human kinase cdna
10	453.2	43.5	2615	22	AAF44670 Novel protein kina
11	251.8	24.2	1698	21	AA29745 Rabbit KTIAMRE kin

DR WPI; 2001-266166/27.  
DR P-PSDB; AAE00494.

PT New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases  
XX  
PS Claim 1; Page 31-32; 38pp; English.

The present sequence is a cDNA encoding novel human protein (NHP) known as human kinase. The human kinases share structural similarity with animal kinases. The human kinases share structural similarity with human kinases, more particularly serine or threonine protein kinases. Human kinase cDNA is useful for the detection of mutant human kinase for the diagnosis of disease, and also as a therapeutic. It is useful for screening drugs, effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying polymorphisms and as primers in amplification assays to detect mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide construct encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a continuous supply of NHP to the body. Nucleotide constructs encoding functional NHPs are used in gene therapy for the modulation of NHP expression.

Sequence 1041 BP; 344 A; 190 C; 228 G; 279 T; 0 other;

Query Match	100.0%;	Score 1041;	DB 22;	Length 1041;
Best Local Similarity	100.0%;	Pred. No. 1.1e-268;		
Matches 1041;	Conservative	0.	Mismatch	

Q:

Qy	1	atggaagaagtataaaaaattagctaagactcgagaaagggtcttatgggtgtgtattcaaa	60
Db	1	atggaagaagtataaaaaattagctaagactcgagaaagggtcttatgggtgtgtattcaaa	60
Qy	61	tgcagaacaaaacctctggacaagtagtagctgttaaaaaatttgtgaatctgaagat	120
Db	61	tgcagaacaaaacctctggacaagtagtagctgttaaaaaatttgtgaatctgaagat	120
Qy	121	gatcctgtgtttaagaaatgacataagagaatacgtatgttgaagcaattataaacat	180
Db	121	gatcctgtgtttaagaaatgacataagagaatacgtatgttgaagcaattataaacat	180
Qy	181	ccaaatcttgaacctcatcgagggtgttcaggagaaaaggaaaatgcatttagttttt	240
Db	181	ccaaatcttgaacctcatcgagggtgttcaggagaaaaggaaaatgcatttagttttt	240
Qy	241	gaatcactgtgatacacactttaaatgagctggaagaagaaacccaaaatggagttgctgat	300
Db	241	gaatcactgtgatacacactttaaatgagctggaagaagaaacccaaaatggagttgctgat	300
Qy	301	ggagtgatacaaaagcgtatattggcaaacaccttcaagctcttaatttctgtcatatacat	360
Db	301	ggagtgatacaaaagcgtatattggcaaacaccttcaagctcttaatttctgtcatatacat	360
Qy	361	aactgtattcacagagataaaaaacctgaaataattctataacttaagcaaggaataatc	420
Db	361	aactgtattcacagagataaaaaacctgaaataattctataacttaagcaaggaataatc	420
Qy	421	aagattttgacttcgggtttgcacaaattctgattccacggagatgcctacacggattat	480
Db	421	aagattttgacttcgggtttgcacaaattctgattccacggagatgcctacacggattat	480
Qy	481	gtagctacagagatggtacacgagctcgtgaactcttctgtggggagatactcagtagtggtct	540

Db	481	gtagctcagagatggtaccagagctcctgaacctctgtggagatactcagtatggtct	540
Qy	541	taadtcgatataatggctattggttgtgtttttgcagagctcctgcagagccagccactg	600
Db	541	taadtcgatataatggctattggttgtgtttttgcagagctcctgcagagccagccactg	600
Qy	601	tggccttgaaaaatcagatgtggaccaactttatctgataatcagaacacatagtagagacg	660
Db	601	tggccttgaaaaatcagatgtggaccaactttatctgataatcagaacacatagtagagacg	660
Qy	661	gggtttgcgcattgtgaccaggtgtgtctgaaactcttgacgtcaagtgatccacctgcc	720
Db	661	gggtttgcgcattgtgaccaggtgtgtctgaaactcttgacgtcaagtgatccacctgcc	720
Qy	721	gtagcctctcaaaagtctggaattacaggaaaaataatcccaagacatacaatacttt	780
Db	721	gtagcctctcaaaagtctggaattacaggaaaaataatcccaagacatacaatacttt	780
Qy	781	aaaagtaaacgggttttccatggcatcagtatacctgagccagaagacatggaactctt	840
Db	781	aaaagtaaacgggttttccatggcatcagtatacctgagccagaagacatggaactctt	840
Qy	841	gaggaagaagtctcagatgttcatcctgtggtctgaaactctcatgaagggtgtctgaag	900
Db	841	gaggaagaagtctcagatgttcatcctgtggtctgaaactctcatgaagggtgtctgaag	900
Qy	901	atgaatccagatgacagattaacctgttcccaactcttgagagctcctactttgattct	960
Db	901	atgaatccagatgacagattaacctgttcccaactcttgagagctcctactttgattct	960
Qy	961	tttcaagagcccaaatataaagaagaagcgcgaatagaagaagaacagagaagcccaa	1020
Db	961	tttcaagagcccaaatataaagaagaagcgcgaatagaagaagaacagagaagcccaa	1020
Qy	1021	caggtaacttcgcgtcaaaagt	1041
Db	1021	caggtaacttcgcgtcaaaagt	1041
RESULT 2			
AAAD03813			
ID	AAAD03813 standard; cdna; 1068 bp.		
AC	AAAD03813;		
DT	19-JUN-2001 (first entry)		
DE	Human kinase cDNA #2.		
KW	Human; kinase; gene therapy; bioreactor; mental disorder;		
KW	biological disorder; ss.		
OS	Homo sapiens.		
Key	Location/Qualifiers		
FT	1..1068		
CDS	/tag= a		
FT	/product= "Human kinase #2"		
FT	/note= "The coding region does not include stop codon"		
FT	/partial		
XX	WO200123579-A1.		
XX	05-APR-2001.		
XX	27-SEP-2000; 2000WO-US26621.		
XX	28-SEP-1999; 99US-0156511.		
PA	(LEXI-) LEXICON GENETICS INC.		
PI	Donoho G, Turner CA, Nehls M, Friedrich G, ZambrowiczB, Sands AT;		

XX WPI: 2001-266166/27.  
DR P-PSDB; AAE00491.  
XX  
XX New isolated human kinase polynucleotide useful for generating for  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases  
XX  
XX Disclosure; Page 28; 38pp; English.  
XX  
XX The present sequence is a cDNA encoding novel human protein  
CC (NHP) known as human kinase. The human kinases share structural  
CC similarity with animal kinases. The human kinases share structural  
CC similarity with animal kinases, more particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.  
XX  
XX Sequence 1068 BP; 344 A; 197 C; 240 G; 287 T; 0 other;

Query Match 98.5%; Score 1025; DB 22; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 2.le-264;  
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggaagaagtataaaattagctgaagctggaaggggtcttatgggtgtattcaaa 60  
Db 1 atggaagaagtataaaattagctgaagctggaaggggtcttatgggtgtattcaaa 60  
Qy 61 tgcagaacaaacctctgacaaagttagctgttaaaaaattgtggaatctgaagat 120  
Db 61 tgcagaacaaacctctgacaaagttagctgttaaaaaattgtggaatctgaagat 120  
Qy 121 gatcctgtgttaagaaatagcactagagaaatacgtatgttgaagaactaaacat 180  
Db 121 gatcctgtgttaagaaatagcactagagaaatacgtatgttgaagaactaaacat 180  
Qy 181 ccaaatctgtgaacctcagctgaggtgttcagagagaaaaagaaaatcatttagtttt 240  
Db 181 ccaaatctgtgaacctcagctgaggtgttcagagagaaaaagaaaatcatttagtttt 240  
Qy 241 gaatactgtgatacaacttttaaatgagctggaagaacccaaatggagttgctgat 300  
Db 241 gaatactgtgatacaacttttaaatgagctggaagaacccaaatggagttgctgat 300  
Qy 301 ggaagtatacaagcgtattatggcaaacacttcaagctcttaatttctgcatacat 360  
Db 301 ggaagtatacaagcgtattatggcaaacacttcaagctcttaatttctgcatacat 360  
Qy 361 aactgtattcacagagataaaaaacctgaaataatttcaataactaagaagaataac 420  
Db 361 aactgtattcacagagataaaaaacctgaaataatttcaataactaagaagaataac 420  
Qy 421 aagatttgtacttcgggtttgcacaaattctgattccaggagatgcctacacgattat 480  
Db 421 aagatttgtacttcgggtttgcacaaattctgattccaggagatgcctacacgattat 480

Qy 481 gtgactacgagatggtaccgagctcctgaactcttctgtggagatactcagtaggttct 540  
Db 481 gtgactacgagatggtaccgagctcctgaactcttctgtggagatactcagtaggttct 540  
Qy 541 tcaatcgatataatggctattggtgtgttttgcagagctcctgacagccagccactg 600  
Db 541 tcaatcgatataatggctattggtgtgttttgcagagctcctgacagccagccactg 600  
Qy 601 tggcctggaataatcgatgtggaccacactttatctgataatcagaacactagtagagacg 660  
Db 601 tggcctggaataatcgatgtggaccacactttatctgataatcagaacactagtagagacg 660  
Qy 661 ggggttcgcctgattgacacagctggtctcgaaactcttgacgtcaagtcatccactgcc 720  
Db 661 ggggttcgcctgattgacacagctggtctcgaaactcttgacgtcaagtcatccactgcc 720  
Qy 721 gtacccctcaaaagtctggaataacaggaataaataatcccaagacatcaatcatttt 780  
Db 721 gtacccctcaaaagtctggaataacaggaataaataatcccaagacatcaatcatttt 780  
Qy 781 aaaaagtaacgggttttccatggcatcagtatcctgagccagacagacatggaaactctt 840  
Db 781 aaaaagtaacgggttttccatggcatcagtatcctgagccagacagacatggaaactctt 840  
Qy 841 gaggaagaagtctcagatgttcatcctgtggtctggaacttcagaggggtgtctgaag 900  
Db 841 gaggaagaagtctcagatgttcatcctgtggtctggaacttcagaggggtgtctgaag 900  
Qy 901 atgaatccagatgacagataaactgttcccaactcctggagagctcctacttatttct 960  
Db 901 atgaatccagatgacagataaactgttcccaactcctggagagctcctacttatttct 960  
Qy 961 ttccaagagcccaataataaaagaaagcagctgaatgaagaaagaaacacagaccccaa 1020  
Db 961 ttccaagagcccaataataaaagaaagcagctgaatgaagaaagaaacacagaccccaa 1020  
Qy 1021 caggt 1025  
Db 1021 caggt 1025  
RESULT 3  
AAD03817  
ID AAD03817 standard; cDNA; 945 BP.  
XX  
AC AAD03817;  
XX  
DT 19-JUN-2001 (first entry)  
XX  
DE Human kinase cDNA #6.  
XX  
KW Human; kinase; gene therapy; bioreactor; mental disorder;  
KW biological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..945  
FT /tag= a  
FT /product= "Human kinase #6"  
FT /note= "The coding region does not include stop codon"  
FT /partial  
XX  
PN WO200123579-A1.  
XX  
PD 05-APR-2001.  
XX  
PF 27-SEP-2000; 2000WO-US26621.  
XX  
PR 28-SEP-1999; 99US-0156511.  
XX (LEXI-) LEXICON GENETICS INC.  
XX



XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
PI WPI: 2001-266166/27.  
DR P-PSDB; AAE00492.  
DR

XX New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases -  
PT  
XX Disclosure; Page 29-30; 38pp; English.

XX The present sequence is a cDNA encoding novel human protein  
XX (NHP) known as human kinase. The human kinases share structural  
CC similarity with animal kinases, more particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs,  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.

XX Sequence 972 BP; 323 A; 172 C; 214 G; 263 T; 0 other;

Query Match 79.1%; Score 823; DB 22; Length 972;  
Best Local Similarity 90.6%; Pred. No. 2e-210;  
Matches 929; Conservative 0; Mismatches 0; Indels 96; Gaps 1;

Qy 1 atggaaaagtataaaaattagctaaagactggagaagggtcttatgggttgattcaaa 60  
Db 1 atggaaaagtataaaaattagctaaagactggagaagggtcttatgggttgattcaaa 60  
Qy 61 tgcagaaacaaacctctgcacaagttagctgttaaaaaatttgtagaactcgaagat 120  
Db 61 tgcagaaacaaacctctgcacaagttagctgttaaaaaatttgtagaactcgaagat 120  
Qy 121 gatcctgttgaagaaatagcactaagagaaatacgtatgttgaagcaattaaacat 180  
Db 121 gatcctgttgaagaaatagcactaagagaaatacgtatgttgaagcaattaaacat 180  
Qy 181 ccaactctgtgaacctcgcaggtgttcagagaaaaagaaaaatcattagtttt 240  
Db 181 ccaactctgtgaacctcgcaggtgttcagagaaaaagaaaaatcattagtttt 240  
Qy 241 gaatactgtatcatatacaacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300  
Db 241 gaatactgtatcatatacaacttttaaatgagctggaaagaaaccccaaatggagttgctgat 300  
Qy 301 ggaatgatacaagcgtattatggaacaccttcaagctcttaattctgtcatatacat 360  
Db 301 ggaatgatacaagcgtattatggaacaccttcaagctcttaattctgtcatatacat 360  
Qy 361 aactgtattcacagagatatataaacctgaaaaatattcttaataactaagcaaggaataatc 420  
Db 361 aactgtattcacagagatatataaacctgaaaaatattcttaataactaagcaaggaataatc 420  
Qy 421 aagatttgacttcgggtttgcacaaattctgattccaggagatgcctacaccgattat 480

Db 421 aagatttgacttcgggtttgcacaaattctgattccaggagatgcctacaccgattat 480  
Qy 481 tgagctacagagatgtaccgagctcctgaactctcttgggagatactcagtagttct 540  
Db 481 tgagctacagagatgtaccgagctcctgaactctcttgggagatactcagtagttct 540  
Qy 541 tcaatcattatagctattggttgggttttttggagagctcctgcagcagccactg 600  
Db 541 tcaatcattatagctattggttgggttttttggagagctcctgcagcagccactg 600  
Qy 601 tggcctggaaaatcagatgtggacccaactttatctgataatcagaacacacacac 660  
Db 601 tggcctggaaaatcagatgtggacccaactttatctgataatcagaacacacacac 660  
Qy 661 ggggttcgcatgttgaccagcgtggtctcgaaactcttgacgtcaagtgatccacctgcc 720  
Db 661 ggggttcgcatgttgaccagcgtggtctcgaaactcttgacgtcaagtgatccacctgcc 720  
Qy 651 ----- 650  
Db 651 ----- 650  
Qy 721 gtgacctctcaagtgctggaattacaggaaaaattatcccaagacatcaatcatttt 780  
Db 721 gtgacctctcaagtgctggaattacaggaaaaattatcccaagacatcaatcatttt 780  
Qy 781 aaaaataacgggttttccatgcatcagtatccctgagccagagacatggaaactctt 840  
Db 781 aaaaataacgggttttccatgcatcagtatccctgagccagagacatggaaactctt 840  
Qy 841 gagggaaaagtctcagatgttccatcctgtggtctctgaaactctgaagggtgtctgaag 900  
Db 841 gagggaaaagtctcagatgttccatcctgtggtctctgaaactctgaagggtgtctgaag 900  
Qy 901 atgaatccagatgacagatataacctgttcccaactcctgagagctcctactttgattct 960  
Db 901 atgaatccagatgacagatataacctgttcccaactcctgagagctcctactttgattct 960  
Qy 961 ttccaaggagcccaaaattaaaaagaaagacgtatgaaggagaaacagagacgcaaa 1020  
Db 961 ttccaaggagcccaaaattaaaaagaaagacgtatgaaggagaaacagagacgcaaa 1020  
Qy 1021 caggt 1025  
Db 1021 caggt 1025  
Qy 925 caggt 929  
Db 925 caggt 929  
RESULT 5  
AAS06725 ID AAS06725 standard; cDNA; 1083 BP.  
AC AAS06725;  
XX  
DT 12-SEP-2001 (first entry)  
XX  
DE Polynucleotide sequence encoding human protein kinase #25.  
XX  
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
KW metabolic disorder; immune related disease; neurological disorder;  
KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
KW reproductive disorder; gene therapy; ss.  
OS Homo sapiens.  
XX  
PN WO200138503-A2.  
XX  
PD 31-MAY-2001.  
XX  
PF 22-NOV-2000; 2000WO-US32085.  
XX  
PR 24-NOV-1999; 99US-0167482.  
XX  
XX (SUGE-) SUGEN INC.  
PA Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
PI Flanagan P, Clary D;  
XX



DR WPI: 2001-266166/27.  
DR P-PSDB; AAE00490.  
XX  
PT New isolated human kinase polynucleotide useful for generating  
PT antibodies, as reagents in diagnostic assays and for screening for  
PT compounds useful for treating mental, biological or medical diseases  
XX  
PS  
PS Disclosure: Page 34-35; 38pp: English.  
XX  
XX The present sequence encodes a novel human protein (NHP) which  
CC is a human kinase. A polymorphism was identified in the 3' UTR  
CC of the present sequence. The human kinases share structural  
CC similarity with animal kinases, more particularly serine or  
CC threonine protein kinases. Human kinase cDNA is useful for the  
CC detection of mutant human kinase for the diagnosis of disease,  
CC and also as a therapeutic. It is useful for screening drugs  
CC effective in the treatment of symptomatic or phenotypic  
CC manifestations perturbing the normal function of NHP in the  
CC body. The NHP nucleotide sequences are useful for generation of  
CC antibodies, as reagents in diagnostic assays, for the  
CC identification of other cellular gene products related to human  
CC kinases, and as reagents in assays for screening compounds that  
CC are useful for treating mental, biological or medical disorders.  
CC NHP oligonucleotides are used as probes. The labelled NHP probes  
CC are useful for screening human genomic library for identifying  
CC polymorphisms and as primers in amplification assays to detect  
CC mutations within the exons, introns and splice sites that can  
CC be used in diagnostics and pharmacogenomics. Nucleotide construct  
CC encoding NHP products are used to genetically engineer cells  
CC in vivo that functions as bioreactors in the body delivering a  
CC continuous supply of NHP to the body. Nucleotide constructs  
CC encoding functional NHPs are used in gene therapy for the  
CC modulation of NHP expression.  
XX  
XX Sequence 1819 BP; 539 A; 368 C; 406 G; 505 T; 1 other;

Query Match 68.1%; Score 709; DB 22; Length 1819;  
Best Local Similarity 77.0%; Pred. No. 7.8e-180;  
Matches 1025; Conservative 0; Mismatches 0; Indels 306; Gaps 1;

QY 1 atggaaagatgataaaattagctaaagactggagaggggtcttattgggtgtattcaaa 60  
DB 138 atggaaagatgataaaattagctaaagactggagaggggtcttattgggtgtattcaaa 197  
QY 61 tgcagaaacaaacctctggacagtagtagctgttaaaaaatttgggaatctgaagat 120  
DB 198 tgcagaaacaaacctctggacagtagtagctgttcaaaaaatttgggaatctgaagat 257  
QY 121 gatcctgttgaagaaatagcactaaagagaaatcgtatgttgaagcaattataaacat 180  
DB 258 gatcctgttgaagaaatagcactaaagagaaatcgtatgttgaagcaattataaacat 317  
QY 181 ccaaatcttgaaacctcatcaggtgttcagagagaaagaaatgcatcttagtttt 240  
DB 318 ccaaatcttgaaacctcatcaggtgttcagagagaaagaaatgcatcttagtttt 377  
QY 241 gaatactgtatcatcaccttttaattgagctggaaagaaacccaaatggagttgtgat 300  
DB 378 gaatactgtatcatcaccttttaattgagctggaaagaaacccaaatggagttgtgat 437  
QY 301 ggagtgatacaaacgctattatggcaaacacttcaagctcttaatttctgcatacatat 360  
DB 438 ggagtgatacaaacgctattatggcaaacacttcaagctcttaatttctgcatacatat 497  
QY 361 aactgtatttcagagatataaaacctgaaatatttctaataactaagaagggaataatc 420  
DB 498 aactgtatttcagagatataaaacctgaaatatttctaataactaagaagggaataatc 557  
QY 421 aagatttgacttcgggtttgcacaaattctg----- 453  
DB 558 aagatttgacttcgggtttgcacaaattcttgagcttcaattcttctctgtgtgcc 617

QY 454 ----- 453  
DB 618 tccattgattggcttaataagttgaccttctgaattcttttcttgcacaaatcagagattttt 677  
QY 454 ----- 453  
DB 678 ctctggttggatccattgctgacacagtgtttccaccatggggcccggtctcatctcga 737  
QY 454 ----- 453  
DB 738 acttggctccaagtgatcttccacctggcctcccaaaagtgtggtgattgcaagtgtg 797  
QY 454 ----- 453  
DB 798 agccacgtgcccagccagatttttcaaacataactactgagagctcacaagattgttt 857  
QY 454 ----- 474  
DB 858 ttatgtgggaacacaatttcgaacaaattcttgagaacgcattccagagagatgcctacacc 917  
QY 475 gattatgtagctacagatgggtaccagctcctgaacttcttctgtgggagatactcagtat 534  
DB 918 gattatgtagctacagatgggtaccagctcctgaacttcttctgtgggagatactcagtat 977  
QY 535 gttcttcagctgatatatggctattgtgtgttttgcagagctcctgcagagccag 594  
DB 978 gttcttcagctgatatatgggtattgtgtgttttgcagagctcctgcagagccag 1037  
QY 595 ccactgtgctggaaaaatcagatgtgtgaccaaactttatctgataatcagaacataagta 654  
DB 1038 ccactgtgctggaaaaatcagatgtgtgaccaaactttatctgataatcagaacataagta 1097  
QY 655 gagacgggttctgcacatgttgaccagctgggtcttcgaactctgcagctcaagtgtacca 714  
DB 1098 gagacgggttctgcacatgttgaccagctgggtcttcgaactctgcagctcaagtgtacca 1157  
QY 715 cctgcctagctctctcaaaagtgtggaattacaggaataattatcccaagacatacaatca 774  
DB 1158 cctgcctagctctctcaaaagtgtggaattacaggaataattatcccaagacatacaatca 1217  
QY 775 atctttaaagataacgggtttttccatggcatcagctacacctgagccagagaacatggaa 834  
DB 1218 atctttaaagataacgggtttttccatggcatcagctacacctgagccagagaacatggaa 1277  
QY 835 actcttgagaaaaagtcttcagatgttccctgtgtgtctggaacttcagaaaggggtgt 894  
DB 1278 actcttgagaaaaagtcttcagatgttccctgtgtgtctggaacttcagaaaggggtgt 1337  
QY 895 ctgaagatgaatccagatgacagattaacctgttcccaactcctggagagctcctacttt 954  
DB 1338 ctgaagatgaatccagatgacagattaacctgttcccaactcctggagagctcctacttt 1397  
QY 955 gattctttcaagaggcccaaatataaagaaaaagcgcgttaattgaaggagaagaaagaga 1014  
DB 1398 gattctttcaagaggcccaaatataaagaaaaagcgcgttaattgaaggagaagaaagaga 1457  
QY 1015 cgcacaacaggt 1025  
DB 1458 cyccaaacaggt 1468

RESULT 7  
AAAF44669  
ID AAF44669 standard; cDNA; 911 BP.  
XX  
AC AAF44669;  
XX  
DT 27-MAR-2001 (first entry)  
DE Novel protein kinase cDNA, SEQ ID NO: 49.  
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;

KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;  
 KW immune disorder; cardiovascular disease; neurodegenerative disease;  
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.  
 XX Homo sapiens.  
 OS  
 XX WO200073469-A2.  
 PN  
 XX 07-DEC-2000.  
 XX  
 XX 26-MAY-2000; 2000WO-US14842.  
 PF  
 XX 28-MAY-1999; 99US-0136503.  
 PR  
 XX (SUGE-) SUGEN INC.  
 PA  
 XX  
 XX Plowman GD, Martinez R, Whyte D, Sudersanam S;  
 PI  
 XX WPI: 2001-032161/04.  
 DR  
 XX P-PSDB; AAB65642.  
 DR  
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
 PT treating immune-related diseases and disorders, cardiovascular disease,  
 PT neurodegenerative diseases and/or cancers -  
 XX  
 XX Example 4; Fig 2; 310pp; English.  
 XX  
 CC The present sequence encodes a novel protein kinase. The nucleic acids  
 CC and the protein kinases they encode may be used in the treatment and  
 CC diagnosis of diseases associated with inappropriate kinase expression  
 CC such as immune-related diseases and disorders, cardiovascular disease,  
 CC neurodegenerative diseases and/or cancers. The nucleic acids and  
 CC complementary sequences may also be used as DNA probes in diagnostic  
 CC assays. The kinase polypeptides may be used as antigens in the production  
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
 CC and kinase antagonists may also be used to down regulate kinase  
 CC expression and activity. Diseases related to kinase expression and  
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
 CC disorders, complications of organ transplantation, myocardial infarction,  
 CC immune disorders, cardiomyopathies, strokes, renal failure,  
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
 CC reproductive disorders.  
 XX  
 XX Sequence 911 BP; 293 A; 182 C; 178 G; 258 T; 0 other;

Query Match 51.6%; Score 537; DB 22; Length 911;  
 Best Local Similarity 87.0%; Pred. No. 5.8e-134;  
 Matches 643; Conservative 0; Mismatches 0; Indels 96; Gaps 1;  
 QY 285 aaatggagtgctgagtgatcaaaagcgtattatggcaaacactcaagctcttaa 344  
 Db 1 aaatggagtgctgagtgatcaaaagcgtattatggcaaacactcaagctcttaa 60  
 QY 345 ttctgtcatatacaatacctgtatttcacagagatataaaacctgaaatatcttaataac 404  
 Db 61 ttctgtcatatacaatacctgtatttcacagagatataaaacctgaaatatcttaataac 120  
 QY 405 taagcaagaataatacaagatttgacttcggttgcaaaattctgattccagaga 464  
 Db 121 taagcaagaataatacaagatttgacttcggttgcaaaattctgattccagaga 180  
 QY 465 tgctacaccgattatgtagctacagatggaccagctcctgaacttctgtggaga 524  
 Db 181 tgctacaccgattatgtagctacagatggaccagctcctgaacttctgtggaga 240  
 QY 525 tactcagtagtggcttctcagtcgatatatggctattgtgtgtttttcagagctct 584  
 Db 241 tactcagtagtggcttctcagtcgatatatggctattgtgtgtttttcagagctct 300

QY 585 gacagccagccactgtgacctggaatacagatgtggaccaactttatctataatcag 644  
 Db 301 gacagccagccactgtgacctggaatacagatgtggaccaactttatctataatcag 360  
 QY 645 aacactagtagagacggggtttcccatgttgaccaggctgtgctcgaactcttgacgtc 704  
 Db 361 aacact----- 366  
 QY 705 aagtgatccaccgtccgtagcctctcaaaagtgtggaattacagaaaaaattaatcccaag 764  
 Db 367 -----agaaaaattaatcccaag 384  
 QY 765 acatcaatacaatctttaaaagtaacgggtttttccatgcgcatacctgagccaga 824  
 Db 385 acatcaatacaatctttaaaagtaacgggtttttccatgcgcatacctgagccaga 444  
 QY 825 agacatggaactcttgaggaaaagtctcagatgttccatcctgtgctcgaacttcat 884  
 Db 445 agacatggaactcttgaggaaaagtctcagatgttccatcctgtgctcgaacttcat 504  
 QY 885 gaagggtgtctgaagatgaatccagatgacagattaacctgttccaaactcctggagag 944  
 Db 505 gaagggtgtctgaagatgaatccagatgacagattaacctgttccaaactcctggagag 564  
 QY 945 ctctactttgattcttttcaagaggcccaaatataaagaaaagcacgtaatgaaggaag 1004  
 Db 565 ctctactttgattcttttcaagaggcccaaatataaagaaaagcacgtaatgaaggaag 624  
 QY 1005 aaacagaagacgcccaacag 1023  
 Db 625 aaacagaagacgcccaacag 643  
 RESULT 8  
 AAD03812  
 ID AAD03812 standard; cDNA; 561 BP.  
 XX  
 AC AAD03812;  
 XX  
 DT 19-JUN-2001 (first entry)  
 XX  
 DE Human kinase cDNA #1.  
 XX  
 KW Human; kinase; gene therapy; bioreactor; mental disorder;  
 KW biological disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH CDS 1..561  
 FT /\*tag= a  
 FT /product= "Human kinase #1"  
 FT /note= "The coding region does not include stop codon"  
 XX /partial  
 PN WO200123579-A1.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 27-SEP-2000; 2000WO-US26621.  
 XX  
 PR 28-SEP-1999; 99US-0156511.  
 XX  
 PA (LEXI-) LEXICON GENETICS INC.  
 XX  
 PT Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;  
 XX  
 DR WPI: 2001-266166/27.  
 DR P-PSDB; AAE00490.  
 XX  
 PT New isolated human kinase polynucleotide useful for generating  
 PT antibodies, as reagents in diagnostic assays and for screening for  
 PT compounds useful for treating mental, biological or medical diseases -



XX	PS	Disclosure; Page 27; 38pp; English.	
XX	XX	The present sequence is a cDNA encoding novel human protein (NHP) known as human kinase. The human kinases share structural similarity with animal kinases, more particularly serine or threonine protein kinases. Human kinase cDNA is useful for the detection of mutant human kinase for the diagnosis of disease, and also as a therapeutic. It is useful for screening drugs effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying polymorphisms and as primers in amplification assays to detect mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide construct encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a continuous supply of NHP to the body. Nucleotide constructs encoding functional NHPs are used in gene therapy for the modulation of NHP expression.	
XX	SQ	Sequence 561 BP; 189 A; 87 C; 112 G; 173 T; 0 other;	
Query Match 43.7%; Score 454.4; DB 22; Length 561;			
Best Local Similarity 99.8%; Pred. No. 5.8e-112;			
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	atggaagaatgataaaattagctaaagactggagaagggtctttaggggtgtattcaaa	60
Db	1	atggaagaatgataaaattagctaaagactggagaagggtctttaggggtgtattcaaa	60
Qy	61	tgcagaacaaacctctggacaagttagctgttaaaaaatttggaatctgaagat	120
Db	61	tgcagaacaaacctctggacaagttagctgttaaaaaatttggaatctgaagat	120
Qy	121	gatcctgtgttaagaaaatagcactaagagaatacgtatgttgaagcaattataaacat	180
Db	121	gatcctgtgttaagaaaatagcactaagagaatacgtatgttgaagcaattataaacat	180
Qy	181	ccaaatcttgaaacctcagctaggtgttcagagaaaagaaatgcatttagtttt	240
Db	181	ccaaatcttgaaacctcaggtgttcagagaaaagaaatgcatttagtttt	240
Qy	241	gaactgtgtatcacacttttaaatgagctggaaagaaacccaaatggagttgctgat	300
Db	241	gaactgtgtatcacacttttaaatgagctggaaagaaacccaaatggagttgctgat	300
Qy	301	ggagtgatcaaaagcgtattatggcaaaccttcaagctcttaattctgcataacat	360
Db	301	ggagtgatcaaaagcgtattatggcaaaccttcaagctcttaattctgcataacat	360
Qy	361	aactgtattcacagagatataaaacctgaaaatttcttaataactaagaaggaataac	420
Db	361	aactgtattcacagagatataaaacctgaaaatttcttaataactaagaaggaataac	420
Qy	421	aagatttgacttcgggtttgcacaaattctgatt	456
Db	421	aagatttgacttcgggtttgcacaaattctgatt	456
RESULT 9			
XX	AAAD03815		
XX	AAAD03815	standard; cDNA; 594 BP.	
XX	AAAD03815;		
XX	XX		

DT	19-JUN-2001	(first entry)	
XX	XX	Human kinase cDNA #4.	
XX	KW	Human; Kinase; gene therapy; bioreactor; mental disorder; biological disorder; ss.	
XX	OS	Homo sapiens.	
XX	FT	Key	Location/Qualifiers
XX	FT	CDS	1..594
XX	FT		/*tag= a
XX	FT		/product= "Human Kinase #4"
XX	FT		/note= "The coding region does not include stop codon"
XX	FT		/partial
XX	XX	WO200123579-A1.	
XX	PD	05-APR-2001.	
XX	PF	27-SEP-2000; 2000WO-US26621.	
XX	PR	28-SEP-1999; 99US-0156511.	
XX	PA	(LEXI-) LEXICON GENETICS INC.	
XX	PI	Donoho G, Turner CA, Nehls M, Friedrich G, zambrowicz B, Sands AT;	
XX	DR	WPI; 2001-266166/27.	
XX	DR	P-PSDB; AAE00493.	
XX	PT	New isolated human kinase polynucleotide useful for generating antibodies, as reagents in diagnostic assays and for screening for compounds useful for treating mental, biological or medical diseases	
XX	PS	Disclosure; Page 31; 38pp; English.	
XX	CC	The present sequence is a cDNA encoding novel human protein (NHP) known as human kinase. The human kinases share structural similarity with animal kinases, more particularly serine or threonine protein kinases. Human kinase cDNA is useful for the detection of mutant human kinase for the diagnosis of disease, and also as a therapeutic. It is useful for screening drugs effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of NHP in the body. The NHP nucleotide sequences are useful for generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products related to human kinases, and as reagents in assays for screening compounds that are useful for treating mental, biological or medical disorders. NHP oligonucleotides are used as probes. The labelled NHP probes are useful for screening human genomic library for identifying polymorphisms and as primers in amplification assays to detect mutations within the exons, introns and splice sites that can be used in diagnostics and pharmacogenomics. Nucleotide construct encoding NHP products are used to genetically engineer cells in vivo that functions as bioreactors in the body delivering a continuous supply of NHP to the body. Nucleotide constructs encoding functional NHPs are used in gene therapy for the modulation of NHP expression.	
XX	SQ	Sequence 594 BP; 203 A; 94 C; 115 G; 182 T; 0 other;	
Query Match 43.7%; Score 454.4; DB 22; Length 594;			
Best Local Similarity 99.8%; Pred. No. 6e-112;			
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	atggaagaatgataaaattagctaaagactggagaagggtctttaggggtgtattcaaa	60
Db	1	atggaagaatgataaaattagctaaagactggagaagggtctttaggggtgtattcaaa	60
Qy	61	tgcagaacaaacctctggaagaatgagctgttataaaatttggaatctgaagat	120

|||||  
Db 61 tgcagaaacaaacctctggacaagtagcgtgttcaaaaaatttggaaatcgaagat 120  
Qy 121 gatcctgtgttaagaaaaatagcactaagagaaatcacgtatgttgaagcaattaaaaaat 180  
Db 121 gatcctgtgttaagaaaaatagcactaagagaaatcacgtatgttgaagcaattaaaaaat 180  
Qy 181 ccaaatcttgtgaacctcatcgaggtgttcagagaaaaagaaaaatgcatttagtttt 240  
Db 181 ccaaatcttgtgaacctcatcgaggtgttcagagaaaaagaaaaatgcatttagtttt 240  
Qy 241 gaactgtgatcacacttttaaatgagctggaagaaaccccaaatggagttctgat 300  
Db 241 gaactgtgatcacacttttaaatgagctggaagaaaccccaaatggagttctgat 300  
Qy 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgcatacat 360  
Db 301 ggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgcatacat 360  
Qy 361 aactgtattcacagagatataaaacctgaaaaatattcttaataactaagcaaggaataatc 420  
Db 361 aactgtattcacagagatataaaacctgaaaaatattcttaataactaagcaaggaataatc 420  
Qy 421 aagatttgtacttcgggtttgcacaaattctgatt 456  
Db 421 aagatttgtacttcgggtttgcacaaattctgatt 456

RESULT 10

AAF44670  
ID AAF44670 standard: cDNA; 2615 BP.  
XX  
AC AAF44670;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE Novel protein kinase cDNA, SEQ ID NO: 50.  
XX  
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;  
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;  
KW immune disorder; cardiovascular disease; neurodegenerative disease;  
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.  
XX  
OS Mus musculus.  
XX  
PN WO200073469-A2.  
XX  
PD 07-DEC-2000.  
XX  
PF 26-MAY-2000; 2000WO-US14842.  
XX  
PR 28-MAY-1999; 99US-0136503.  
XX  
PA (SUGEN-) SUGEN INC.  
XX  
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;  
DR WPI; 2001-032161/04.  
DR P-PSDB; AAB65643.  
XX  
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
PT treating immune-related diseases and disorders, cardiovascular disease,  
PT neurodegenerative diseases and/or cancers -  
XX  
PS Disclosure; Fig 2; 310pp; English.  
XX  
CC The present sequence encodes a novel protein kinase. The nucleic acids  
CC and the protein kinases they encode may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and

complementary sequences may also be used as DNA probes in diagnostic assays. The kinase polypeptides may be used as antigens in the production of antibodies of kinase expression and activity. Anti-kinase antibodies and kinase antagonists may also be used to down regulate kinase expression and activity. Diseases related to kinase expression and activity include rheumatoid arthritis, atherosclerosis, autoimmune disorders, complications of organ transplantation, myocardial infarction, immune disorders, cardiomyopathies, strokes, renal failure, oxidative-stress related disorders, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive disorders.  
XX  
SQ Sequence 2615 BP; 742 A; 580 C; 585 G; 708 T; 0 other;  
  
Query Match 43.5%; Score 453.2; DB 22; Length 2615;  
Best Local Similarity 73.1%; Pred. No. 2.2e-111;  
Matches 658; Conservative 0; Mismatches 143; Indels 99; Gaps 2;  
  
Qy 135 gaaaatagcacaaagaaatacgtatgttgaagcaattaaacatccaaatccttgtaa 194  
Db 1 gaaaatagccctcgccgggaaatccgtatgctgaag---ttgaaacacccaaacctgtaa 57  
  
Qy 195 cctcatcgaggtgttcagagaaaaagaaaatgcatttagtttttgaatactgtgatca 254  
Db 58 cctcatcgaggtgttcagagaaaaagaaagatgcatttagtttttgaatactgtgatca 117  
  
Qy 255 tacacttttaaatgagctggaagaaaccccaaatggaggtgttcagaggaatcaaaag 314  
Db 118 cacactgttaaacagagctggagagaaaccccaacggaggttctgtgagtgattaaaag 177  
  
Qy 315 cgtattatggcaaacacttcaagctcttaatttctgcatacatataactgtattcacag 374  
Db 178 tgtgctatggcaaaccttcaagccttaacttctgtcaacagcaaatgtattcacag 237  
  
Qy 375 agatataaaacctgaaatattctaaactaactaagcaaggaataatcaagatttgactt 434  
Db 238 ggaatgaaacctgaaacatctctaaacaaaggaaggaataagatttgactt 297  
  
Qy 435 cgggtttcacaaattctgtccagagagatgctcacaccgattatgtactacagatg 494  
Db 298 tggatttgcaggaattctaatccagagagcctcacagactatgttgcacacaggtg 357  
  
Qy 495 gtaccgagctctgaacttctgtggagataactcagtagtcttcagtcgatatatg 554  
Db 358 gtaccgagccccgaacttctcgtggagacacgaagtaacgttctctgtagacgtgtg 417  
  
Qy 555 ggcctattgtgtgttttgcagagctcctgaagccagccactgtgacctggaaaaate 614  
Db 418 ggcctgagctgtgttttgcagagctcctgaagccactgtgacctggaaaaate 477  
  
Qy 615 agatgtgacccaactttatctgataatcagaaacactagtagagacggggttgcacatgt 674  
Db 478 cgacgtggaccagcttcaactgatcatcaggacgttg----- 514  
  
Qy 675 tgaccagctggtctcgaacttgaactgaagtcaagtgtatccactgcgtagcctctcaag 734  
Db 515 ----- 514  
  
Qy 735 tgcgtgaattacaggaataatcccaagacatcaatcaatctttaaagtaacgggtt 794  
Db 515 -----gggaagctgattccaaagacaccagctctatcttttaggtaacacagt 561  
  
Qy 795 ttccatggcatcagtagtaccctgaaccagaagacatggaacctttgaggaaaagtcttc 854  
Db 562 ttcccgccgacatcagctacccgaacagggacatggagactcttgaagaaaaattctc 621  
  
Qy 855 aagatttcaactgtgtgcttgaacttcatgaaggggtgtctgaagatgaatccagatga 914  
Db 622 aagatttcaactgtgtgcttgaacttcatgaaggggtgtctgaagatgaatccagatga 681  
  
Qy 915 cagattaaactgttcccaactcctggagagctcctacttcttattcttcaagagccca 974

Db	682	gagcctgacccgtgcccagctgctggacagtgctacttgagctctttcaagagaatca	741
Qy	975	aattaaagaagaacgtaatgaagggaagaacagagacgcaacagctacttccgct	1034
Db	742	aatgaagaagaagcccgagtgagggagaagccgaagcgccagcagaatcaactgct	801
RESULT 11			
ID	AAA29746	standard; DNA; 1698 BP.	
XX	AAA29746:		
DT	15-AUG-2000	(first entry)	
DE	Rabbit KKIAMRE kinase genomic DNA SEQ ID NO:5.		
XX	Rabbit; KKIAMRE kinase; learning-induced kinase; learning; memory;		
KW	cd2-related kinase; brain; gene therapy; genetic disorder; detection;		
KW	identification; ds.		
OS	Oryctolagus cuniculus.		
XX	WO2000020567-A2..		
PN	13-APR-2000.		
PD	01-OCT-1999; 99WO-US23010.		
PF	02-OCT-1998; 98US-0102906.		
PR	(UYS-) UNIV SOUTHERN CALIFORNIA.		
PA	Thompson RF, Gomi H, Sun W;		
PI	WPI: 2000-328932/28.		
DR	Novel learning induced kinase polynucleotides and polypeptides, useful		
XX	for the analysis of learning and memory, and for gene therapy -		
PT	Claim 1; Fig 9; 64pp: English.		
PS	The present sequence encodes a learning-induced kinase, designated		
XX	KKIAMRE kinase, which is isolated from rabbit brain tissue. KKIAMRE		
CC	kinase is a cd2-related kinase. The KKIAMRE kinase polynucleotides		
CC	be used to express recombinant protein for analysis, characterisation		
CC	therapeutic use, as markers for tissues in which the protein is		
CC	preferentially expressed, as molecular weight markers on Southern gel		
CC	as chromosome markers or tags, to compare endogenous DNA sequences in		
CC	patients to identify potential genetic disorders, as probes to hybrid		
CC	and discover novel related sequences, as a source of PCR primers, and		
CC	an antigen to induce anti-DNA antibodies. The polypeptides can be used		
CC	in assay to discover biological activity, to raise antibodies, as tis		
CC	markers, and to isolate correlative receptors or ligands. The		
CC	polynucleotides may also be used for gene therapy for the treatment o		
CC	disorders which are mediated by KKIAMRE kinase.		
XX	Sequence 1698 BP: 551 A; 311 C; 390 G; 446 T; 0 other;		
SQ	Query Match 24.2%; Score 251.8; DB 21: Length 1698;		
	Best Local Similarity 62.9%; Pred. No. 1.3e-57;		
	Matches 407; Conservative 0; Mismatches 237; Indels 3; Gaps		
Qy	1 atggaagaatgaaacacacgtaagactggagaaggcctttaggggttgattcaaa	60	
Db	1 atggaagaatgaaacacacgtaagactggagaaggcctttaggggttgattgaag	60	
Qy	61 tgcagaacacacacacacgtaagactggagaaggcctttaggggttgattgaag	120	
Db	61 tgcagaacacacacacacgtaagactggagaaggcctttaggggttgattgaag	120	

PT Novel learning induced kinase polynucleotides and polypeptides, useful  
XX for the analysis of learning and memory, and for gene therapy -  
PS Claim 1; Fig 4; 64pp; English.

CC The present sequence encodes a learning-induced kinase, designated  
CC KKIAME kinase, which is isolated from rabbit brain tissue. KKIAME  
CC kinase is a cdc2-related kinase. The KKIAME kinase polynucleotides can  
CC be used to express recombinant protein for analysis, characterisation or  
CC therapeutic use, as markers for tissues in which the protein is  
CC preferentially expressed, as molecular weight markers on Southern gels,  
CC as chromosome markers or tags, to compare endogenous DNA sequences in  
CC patients to identify potential genetic disorders, as probes to hybridise  
CC and discover novel related sequences, as a source of PCR primers, and as  
CC an antigen to induce anti-DNA antibodies. The polypeptides can be used  
CC in assay to discover biological activity, to raise antibodies, as tissue  
CC markers, and to isolate correlative receptors or ligands. The  
CC polynucleotides may also be used for gene therapy for the treatment of  
CC disorders which are mediated by KKIAME kinase.

XX Sequence 3080 BP; 906 A; 569 C; 765 G; 840 T; 0 other;

Query Match 24.2%; Score 251.8; DB 21; Length 3080;  
Best Local Similarity 62.9%; Pred. No. 1.6e-57;  
Matches 407; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

QY 1 atggaaaagtataaaaattgctaagactggagaagggtcttattgggtgtgtattcaaa 60  
Db atggaaaattatgagaaccttggtggtggagaagggtattggaatggtgatgaag 308  
QY 61 tgcgagaacaaacctctggcaagtagtagctgttaaaaaatttggaaactctgaagat 120  
Db 309 tgtaggaataagatggaagtgtggccatcaagaagttctctagaaagtgtgat 368  
QY 121 gatcctgttttaagaaaatagcactaagaagaatacgtatgttggaagcaattaaacat 180  
Db 369 gacaaaatggtcaaaaaattgctatgcgagaatacaagttaactaaagcaactgagcat 428  
QY 181 ccaaatctgtgaacctcatcgaggtgttcaggagaaagaaagaaatgattgttttt 240  
Db 429 gaaaattggtgaatctgttgaggtgtgtaaaaaaacacagtggtacctggtgttc 488  
QY 241 gaactactgtgatacacttttaaatgagctggaagaaccccaatggagttgctgat 300  
Db 489 gaattgttgaccacagactcttgatgacttggaaactcttccaaatggactagatgac 548  
QY 301 ggaagtatcaaaagcgtattatggcaaacacttcaagctcttaattctgtcatatacat 360  
Db 549 caagtgttcaaaagtattgtttcagatttaataatggaatggatttggccaagtcac 608  
QY 361 aactgtattcacagataataaacctgaaatattcttaactaagaagaataatc 420  
Db 609 aatatcatatagatataaaagccagagataatattggtctccacagctcgtggtgttc 668  
QY 421 aagattgtgactcgtggtttgcacaaattctgt---attccagagatgctcacacgat 477  
Db 669 aagttatgtattgtgatttgcagcactggcagctcccgagaggtttacactgat 728  
QY 478 tatgtactgacagatggttacagactcctgaactcttctgtggagatactcagttatggt 537  
Db 729 tatgtgcaactcagatggttacagactcctgaagactcgtggtgtgtcaagtatggc 788  
QY 538 tctcaagtcgatatatgggctattgtgtgttttgcagagctcctgcagagccagcca 597  
Db 789 aaagctgtggatgtgtggccattgtgtctgtaactgaaatgctcaatcgtggggaaacc 848  
QY 598 ctgtgctctggaataatcagatgtgaccaacttctatctataatcag 644  
Db 849 ctgttctctggagactctgatattgatcagctttatctctattatgat 895

RESULT 13

AAZ51208  
ID AAZ51208 standard; cDNA; 1513 BP.

XX AC AAZ51208;

XX DT 06-JUN-2000 (first entry)

XX DE Human lost in leukaemia kinase (LLK) cDNA.

XX KW Lost in leukaemia kinase; LLK; cdc2-related kinase; human; leukaemia;  
XX MAPK; mitogen activated protein kinase; tumour suppressor; diagnosis;  
XX chromosome 5q31.1; cytostatic; mitotic index; treatment; prevention;  
XX gene therapy; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 1..1368

XX FT /\*tag= a

XX FT /product= "Human lost in leukaemia kinase protein"

XX PN W0200012719-A1.

XX PD 09-MAR-2000.

XX PF 31-AUG-1999; 99WO-CA00794.

XX PR 31-AUG-1998; 98CA-2243784.

XX PR 20-NOV-1998; 98CA-2251249.

XX PA (ONTA-) ONTARIO CANCER INST.

XX PI Zanke B, Haq R, Randall S, Midmer M;

XX DR WPI: 2000-237880/20.

XX DR P-PSDB; AAY70126.

XX PT Isolated polynucleotide encoding a lost in leukemia kinase (LLK) protein, useful for treatment, diagnosis and prevention of leukemia -

XX PS Claim 4c; Page 51; 69pp; English.

XX CC The present sequence is the cDNA encoding human lost in leukaemia kinase (LLK), isolated from human lambda gt10 foetal heart genomic library. This gene is mapped to chromosome 5q31.1, a region implicated in human acute leukaemia. It is expressed strongly in muscle, heart, liver, brain and in tissues which have a very low mitotic index. Human LLK has 83% amino acid sequence homology to rat LLK beta protein. LLK is closely related to cdc2-related kinases, that are putative tumour suppressors and to mitogen activated protein kinases (MAPKs). This sequence has cytostatic activity. It is useful for the treatment, diagnosis and prevention of acute leukaemia and is also used in gene therapy.

XX SQ Sequence 1513 BP; 551 A; 253 C; 295 G; 414 T; 0 other;

Query Match 19.5%; Score 202.6; DB 21; Length 1513;  
Best Local Similarity 59.4%; Pred. No. 1.8e-44;  
Matches 381; Conservative 0; Mismatches 254; Indels 6; Gaps 2;

QY 1 atggaaaagtataaaaattgctaagactggagaagggtcttattgggtgtgtattcaaa 60  
Db 1 atggagaatgataaaaccttggaaaagtgggagagggaagttaacggaacagtcagtaa 60  
QY 61 tgcgagaacaaacctctggcaagtagtagctgttataaaatttggaaactctgaagat 120  
Db 61 tgaatacataagaataactgtggcagatagtcgcatttaagatatattttagagagaccagaa 120  
QY 121 gatcctgttttaagaaaatagcactaagaagaataacgtatgttggaagcaattaaacat 180  
Db 121 caactctgtc---aacaaaattgcgatgagagaaaataaagtttctctaaagcaatttcac 177

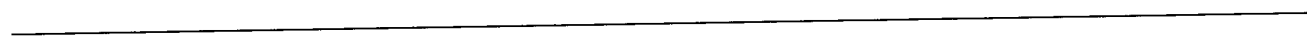




.....

Fri May 3 11:02:13 2002

us-09-671-050-9.rng







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 14:02:28 : Search time 2235.62 Seconds  
(without alignments)  
5003.693 Million cell updates/sec

Title: US-09-671-050-9  
Perfect score: 1041  
Sequence: 1 atggaagatgataaaatt.....aggtaacttcgcgtcaaaagt 1041

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estfun:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estom:\*
  - 5: em\_estpl:\*
  - 6: em\_estba:\*
  - 7: em\_estro:\*
  - 8: em\_estov:\*
  - 9: em\_hic:\*
  - 10: gb\_est1:\*
  - 11: gb\_est2:\*
  - 12: gb\_hic:\*
  - 13: gb\_gss:\*
  - 14: em\_gss\_fun:\*
  - 15: em\_gss\_hum:\*
  - 16: em\_gss\_inv:\*
  - 17: em\_gss\_pln:\*
  - 18: em\_gss\_pro:\*
  - 19: em\_gss\_rtd:\*
  - 20: em\_gss\_vrt:\*
  - 21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	333.6	32.0	500	10	AI385966	ml34h09.y
2	331	31.8	1691	12	AK016781	AK016781 Mus muscu
3	316.8	30.4	938	10	AA061797	AA061797 ml34h09.y
4	258.6	24.8	731	10	AA106692	AA106692 um32d03.y
5	255	24.5	718	10	AA286088	AA286088 vc33b05.y
6	209.8	20.2	579	10	AA233105	AA233105 fj28b10.y
7	199	19.1	295	10	AA626859	AA626859 zu89f10.s
8	186.6	17.9	727	10	AI508835	AI508835 vc33b05.y
9	175.6	16.9	1535	12	BC010966	BC010966 Homo sapi
10	174.8	16.8	582	10	AV672492	AV672492 AV672492
11	174.4	16.8	665	13	AQ953706	AQ953706 Sheared D
12	158.4	15.2	495	10	AI385990	AI385990 ml50h06.y

13	154.6	14.9	744	11	BF787339	BF787339
14	152.6	14.7	879	11	BC281841	BC281841
15	152.6	14.7	926	10	AL524036	AL524036
16	150.4	14.4	779	11	BF760621	BF760621
17	149.6	14.4	384	10	BE223187	BE223187
18	141.4	13.6	808	10	AI746896	AI746896
19	141.2	13.6	795	10	BE275968	BE275968
20	140	13.4	957	11	BI410031	BI410031
21	135.2	13.0	699	10	AL584951	AL584951
22	134.8	12.9	754	11	BF578758	BF578758
23	134.6	12.9	1131	10	BE563495	BE563495
24	134	12.9	202	10	BE135229	BE135229
25	134	12.9	696	11	BF479244	BF479244
26	133.8	12.9	752	10	BE746563	BE746563
27	133.8	12.9	911	11	BG397188	BG397188
28	133	12.8	680	10	AL585015	AL585015
29	133	12.8	1112	10	AT002436	AT002436
30	132.8	12.8	619	11	BE844514	BE844514
31	132.4	12.7	764	11	BG162489	BG162489
32	132.4	12.7	1000	11	BF316057	BF316057
33	131.2	12.6	831	11	BG825471	BG825471
34	131	12.6	297	10	AI965145	AI965145
35	131	12.6	963	11	BG420755	BG420755
36	130.4	12.5	613	10	AL504123	AL504123
37	130.4	12.5	778	10	BE413170	BE413170
38	130.2	12.5	793	11	BG700342	BG700342
39	130.2	12.5	796	11	BG915878	BG915878
40	130	12.5	713	10	AI246661	AI246661
41	129.8	12.5	379	10	AI094352	AI094352
42	129.8	12.5	884	11	BG440417	BG440417
43	129.8	12.5	945	11	BG481522	BG481522
44	129.6	12.4	673	11	BG976698	BG976698
45	129.2	12.4	732	10	AW701949	AW701949

ALIGNMENTS

RESULT 1

AI385966 500 bp mRNA EST 27-JAN-1999  
LOCUS ml34h09.y1 Stratagene mouse testis (#937308) Mus musculus cDNA  
DEFINITION clone IMAGE:513953 5', similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE KKIALRE (HUMAN);, mRNA sequence.

ACCESSION AI385966  
VERSION AI385966  
KEYWORDS house mouse.  
SOURCE EST.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 500)  
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schuck,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
The WashU-NCI Mouse EST Project 1999  
Unpublished (1999)

TITLE The WashU-NCI Mouse EST Project 1999  
JOURNAL  
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:307801  
This read is a RESEQUENCE of a previously sequenced mouse clone  
correct orientation  
Seq primer: -40RP from Gibco  
High quality sequence stop: 339.

FEATURES		Location/Qualifiers	
SOURCE	1. .500	/organism="Mus musculus"	
		/strain="Inbred CD-1"	
		/db_xref="taxon:10090"	
		/clone_image="513953"	
		/clone_lib="Stratagene mouse testis (#937308)"	
		/sex="males"	
		/tissue_type="testis"	
		/dev_stage="10-12 week old"	
		/lab_host="SOLR (kanamycin resistant)"	
		/note="Organ: testis; Vector: pBluescript SK-; Site1: EcoRI; Site2: XhoI; Cloned unidirectionally. Primer: Oligo df. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'"	
BASE COUNT	145 a 114 c 121 g 120 t		
ORIGIN			
Query Match 32.0%; Score 333.6; DB 10; Length 500;			
Best Local Similarity 82.3%; Pred. No. 9.6e-65;			
Matches 408; Conservative 0; Mismatches 84; Indels 4; Gaps 2;			
QY	135	gaaatagcactaagagaaatcgtatgttgagcaattaaacatccaaatctgtgaa	194
Db	2	GAATAAGCCCTCGGGAATCCGTATGCTGAG--TTGAACACCCCAACCTCGTAA	58
QY	195	cctcatcgaggtgttcaggagaaagaaaaatgcatttagtttgaatactgtgac	254
Db	59	CTCATCGAGGTGTTCAAGAAAGAGAAAGATGCTATGTTTGGTACTGTGATCA	118
QY	255	tacacttttaaatgagctggaaagaaacccaaatggatgtctgagtgatcaaaag	314
Db	119	CACACTGTTAAAGAGCTGGAGAAACCAACCGAGTTCGTGAGTGAATTAAG	178
QY	315	cgattatgcaaaacactcaagctcttaatttctgtatcatatcacatcgtattcaag	374
Db	179	TGCTGTATGCAACCCCTCAAGCCCTTACTCTGTGTACCAAGCAATGTTGATCG	238
QY	375	agataaaaaacctgaaatatcttaactaagcaaggaaatcaagatttgtgactt	434
Db	239	GGATGTAACCTGAAACATCTTAATTAACCAAGCAAGGATGATTAAGATTGTGACTT	298
QY	435	cgggttgacaaaattctgtatccagagatcctacacccgattatgttagctacgattg	494
Db	299	TGGATTGCAACGAATTCATTCAGSAGACGCCTACACAGACTATGTTGCCACAGGTG	358
QY	495	gtaccagctctgaactcttggggagatactcagtatggttcttcagtcgatatg	554
Db	359	GTACCGAGCCCTCGAACCTCTCGTGGGAGACACGAAGTACGGTCTCTGTAGACGTGTG	418
QY	555	ggctattggtgtgtttgttcagagctcctgacagggccagccactgtgacctgaaaaatc	614
Db	419	GGCCGTCGGCTGTGTTATGCAGAGCTTCTGAGCTGTACCCACTGCGCCCG-AAAA	477
QY	615	agatgtggaccaactt 630	
Db	478	CGAGCTGGACCACTT 493	
RESULT 2			
AK016781			
LOCUS			
DEFINITION	AK016781	1691 bp mRNA HTC	05-JUL-2001
ACCESSION		Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933411017, full insert sequence.	
VERSION	AK016781.1	GI:12855701	
KEYWORDS		CAP trapper.	
SOURCE		Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library, clone:4933411017.	
ORGANISM		Mus musculus	
FEATURES			
Location/Qualifiers			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 1691)			
Carninci, P. and Hayashizaki, Y.			
High-efficiency full-length cDNA cloning			
Methods in enzymology. 303, 19-44 (1999)			
99279253			
10349636			
2 (bases 1 to 1691)			
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
Genome research. 10 (10), 1617-1630 (2000)			
20499374			
11042159			
3 (bases 1 to 1691)			
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
RIKEN integrated sequence analysis (RISA) system--384-format			
sequencing pipeline with 384 multicapillary sequencer			
Genome research. 10 (11), 1757-1771 (2000)			
20530913			
11076861			
4 (bases 1 to 1691)			
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
Functional annotation of a full-length mouse cDNA collection			
Nature 409, 685-690 (2001)			
5 (bases 1 to 1691)			
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hamaoka, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.			
Direct Submission			
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)			
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.			
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGGATCCAGAGCTCTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGAGTTAATAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised between 0.5 and 3 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.			
Location/Qualifiers			

[illegible]





```

QY 241 gaactgctgaatcatttttaaatgaagctggaagaaccccaaatggadtgtgat 300
    || || || || || || || || || || || || || || || || || || || ||
Db 455 GAGTCTGAGACACACACTGCTGAATGAGCTGGACAGATACCCACGAGGTGTCCAGAG 514
    || || || || || || || || || || || || || || || || || || || ||
QY 301 ggaagtcaaaagcgtattatggcaaacacttcaagctctcttaattctgtcatatacat 360
    || || || || || || || || || || || || || || || || || || || ||
Db 515 CATATGGTTAAAGCATCATCTGGGCAACACTTTAGGCTGTGAACTTCGCCACAACAA 574
    || || || || || || || || || || || || || || || || || || || ||
QY 361 aactg 365
    || ||
Db 575 AATTG 579
    || ||

RESULT 7
AA626859/LOCUS
DEFINITION AA626859 295 bp mRNA EST 15-OCT-1997
            zu89f10.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:745195
            similar to SW:KKIA_HUMAN Q00532 SERINE/THREONINE-PROTEIN KINASE
            KKIALRE ;, mRNA sequence.
ACCESSION AA626859
VERSION AA626859.1 GI:2539246
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 295)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marr,M., Martin
            J., Moore,B., Scheellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,I., Wylie,T., Waterston,R. and Wilson,R.
            WashU-NCI human EST Project
            Unpublished (1997)
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40ml3 fwd. ET from Amersham
            High quality sequence stop: 292.
FEATURES
            source
            1..295
            /organism="Homo sapiens"
            /db_xref="GDB:5933055"
            /db_xref="taxon:9606"
            /clone="IMAGE:745195"
            /clone_lib="Soares_testis_NHT"
            /sex="male"
            /lab_host="DH10B"
            /note="Vector: pMT3D-Pac (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
            was prepared from mRNA obtained from Clontech Laboratories
            , inc., and primed with a Not I - oligo(dT) primer [5',
            TGTTCACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
            Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT73 vector. Library
            went through one round of normalization to Cot5, and was
            constructed by Bento Soares and M. Fatima Bonaldo."
            95 a 69 c 55 g 76 t
BASE COUNT 95 a 69 c 55 g 76 t
ORIGIN

Query Match 19.1%; Score 199; DB 10; Length 295;
Best Local Similarity 95.3%; Pred. No. 1.4e-34;
Matches 205; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 438 gtttcacaaattctgtccagagagatgctacaccgattatgtacgacagatgta 497
    || || || || || || || || || || || || || || || || || || || ||
Db 280 GATTGATGTCATTGCGATTCACGAGATGCTACCCGATTATGTACGAGATGGTA 221
    || || || || || || || || || || || || || || || || || || || ||

```

```

QY 498 ccagctcctgaactctctgtggagaaataactcagtatggttcttcagtcgatatatgggc 557
    || || || || || || || || || || || || || || || || || || || ||
Db 220 CCAGAGTCTCGAAGTCTCTTGTGGGAGATACTCAGTATAGGTCTTCAGTCGATATGGGC 161
    || || || || || || || || || || || || || || || || || || || ||
QY 558 tatgtgtgtgttttgcagagctcctgacagggccagccactgtggcctggaaaatcaga 617
    || || || || || || || || || || || || || || || || || || || ||
Db 160 TATTGGTTGTGTTTTTGAGAGCTCTGACAGGCGCAGCCACTGTGGCCCTGGAAATCAGA 101
    || || || || || || || || || || || || || || || || || || || ||
QY 618 tgggaccaacttatctgataatacagaacactag 652
    || || || || || || || || || || || || || || || || || || || ||
Db 100 TGTGGACCACTTTATCTGATATATCAGAACTAG 66
    || || || || || || || || || || || || || || || || || || || ||

RESULT 8
AI508835/LOCUS
DEFINITION AI508835 727 bp mRNA EST 12-MAR-1999
            vc33b05.y1 Barstead MPLRBI Mus musculus cDNA clone IMAGE:776337 5'
            similar to gb:X66358 SERINE/THREONINE-PROTEIN KINASE KKIALRE (HUMAN
            );, mRNA sequence.
ACCESSION AI508835
VERSION AI508835.1 GI:4407740
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 727)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
            The WashU-NCI Mouse EST Project 1999
            Unpublished (1999)
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:469193
            This read is a RESEQUENCE of a previously sequenced mouse clone
            correct orientation)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 399.
FEATURES
            Location/Qualifiers
            1..727
            /organism="Mus musculus"
            /strain="BALB/c"
            /db_xref="taxon:10090"
            /clone="IMAGE:776337"
            /clone_lib="Barstead MPLRBI"
            /sex="mixed"
            /tissue_type="Kidney"
            /dev_stage="6 weeks"
            /lab_host="DH10B"
            /note="Vector: pT73D-Pac (Pharmacia) with a modified
            polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
            was primed with a Not I - oligo(dT) primer [5',
            TGTWACGAATCTGAGTGGAGCGCGCCGCTTTTTTTTTTTTTTTTTTTTTTTTTTTT
            3']; double-stranded cDNA was ligated to Eco RI adaptors
            [CATGGATTCGGTACC], digested with Not I and cloned into the
            Not I and Eco RI sites of the modified pT73 vector.
            Library constructed by Bob Barstead."
            200 a 169 c 191 g 158 t 9 others
BASE COUNT 200 a 169 c 191 g 158 t 9 others
ORIGIN

```

Query Match 17.9%; Score 186.6; DB 10; Length 727;

Best Local Similarity 64.5%; Pred: NO. 8e-32;			
Matches 287; Conservative 0; Mismatches 157; Indels 1; Gaps 1;			
QY	1	atggaagattgaataatagctaaagactggaagggcttctatgggtgtattcaaaa	60
Db	110	ATGGAAAATATGAAAATTTGGAAAGATTGGAGAAGCTCCTATGGGTAGTGTCAAG	169
QY	61	tgcagaaacaaacccctcgagcaagtagtctgttataaaattttggaatcgaagat	120
Db	170	TGCAGAAACAGGACACGGGTCAGATCGTGGCCATCAAGAGGTTTCTGGAAACCGAGAT	229
QY	121	gatcctgtgttaagaataatagcactaaagaaatcacgtatgttgaagcaatataaacat	180
Db	230	GACCTGTCTATAAGAAAATCGCCCTTCGAGAAATCCGCATGTCTCAAGCACTCAAGCAG	289
QY	181	ccaaatcttggaaacctcatcgaggtgttcagagaaagaaagaaatgcatttaatttt	240
Db	290	CCCAACCTGTCAACCTCTCTGGAAGTCCTTCGGAGGAAGCGGAGGCTTACCTTGGTGTTC	349
QY	241	gaatactgtgtatcacatcaccttttaaatgagctggaagaaaccccaatggagtctgat	300
Db	350	GAGTACTGCCACACACGCTGCTTCACAGCAGTGGATCGGTATCAAGGGGGGTACACAGAG	409
QY	301	ggagtgtcaaaagcgtattatggcaaacacacctcaagctcttaatttctgtcatatacat	360
Db	410	CTNCTCGTGAAGA-CATNNACTTGCAGACACTGCAGGCTGTNAATTTCTGGCATTAACA	468
QY	361	aactgtattcacagagatataaaacctgaaatattctaaactaaagcaaggaataatc	420
Db	469	TACTGCATACCAAGGACCTGAGCGGAAAATATCTCATCACAAAACAGTCNAGCATT	528
QY	421	aagattgtgactctgggtgtgcac	445
Db	529	AAGCCCTGTATNNTTTGGGTTCCGAC	553
RESULT 9			
LOCUS	BC010966	1535 bp	mRNA
DEFINITION	Homo sapiens, Similar to serine/threonine kinase 9, clone		
ACCESSION	IMAGE:4294753, mRNA.		
VERSION	BC010966		
KEYWORDS	BC010966.1 GI:15012116		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 1535)		
JOURNAL	Strausberg, R.		
	Direct Submission		
	Submitted (23-JUL-2001) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
COMMENT	Contact: MGC help desk		
	Email: <a href="mailto:cgapbs-femail.nih.gov">cgapbs-femail.nih.gov</a>		
	Tissue Procurement: CLONTECH		
	cDNA Library Preparation: CLONTECH Laboratories, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Institute for Systems Biology		
	<a href="http://www.systemsbio.org">http://www.systemsbio.org</a>		
	contact: <a href="mailto:amadan@systemsbiology.org">amadan@systemsbiology.org</a>		
	Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia		
	Greene, Mark Kettman and Anuradha Madan		
Clone distribution: MGC clone distribution information can be found			
through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>			
Series: IRAL Plate: 19 Row: 0 Column: 9			
This clone was selected for full length sequencing because it			
passed the following selection criteria: matched mRNA gi: 3559924			
This clone has the following problem: incomplete processing.			

FEATURES		Location/Qualifiers	
source		1..1535	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:4294753"	
		/tissue_type="Prostate"	
		/clone_lib="NIH_MGC_83"	
		/lab_host="DH10B"	
		/note="Vector: pDNR-LIB"	
BASE COUNT		467 a 322 c 361 g 385 t	
ORIGIN			
Query Match		16.9%; Score 175.6; DB 12; Length 1535;	
Best Local Similarity		56.8%; Pred. No. 2.3e-29;	
Matches		366; Conservative 0; Mismatches 269; Indels 9; Gaps 2;	
QY	1	atggaagattgaataatagctaaagactggaagggcttctatgggtgtattcaaaa	60
Db	306	ATGAATAAATTTGAGATCCTTTGGGGTTGTAGGTGAAGAGGCCATATGGAGTTGTATTAA	365
QY	61	tgcagaaacaaacccctcgagcaagtagtctgttataaaatttttggaaatcctgaagat	120
Db	366	TGCAGACACAGGAACACATGAATTTGTGGCGATCAAGAAATTCGAAGACAGTGAAGAA	425
QY	121	gatcctgtgttaagaataatagcactaaagaaatcacgtatgttgaagcaatataaacat	180
Db	426	AATGAAGAAGTCAAGAAGACGACTTTACGAGAGCTTTAAATGCTTCGGACTCTCAAGCAG	485
QY	181	ccaaatcttggaaacctcatcgaggtgttcagagaaagaaagaaatgcatttaatttt	240
Db	486	GAAAACATTTGGAGTTGAAGGAAGCATTTTCGTCGGAGGGGAAGTTGTACTTGGTGT	545
QY	241	gaatactgtgtatcacatcaccttttaaatgagctggaagaaaccccaaatgaatgctgat	300
Db	546	GAGTATGTTGAAAAAATATGCTCGAATTCCTGGAGAATAATGCCAAATGAGGTTCACCT	605
QY	301	ggagtgtcaaaagcgtattatggcaaacacacctcaagctcttaatttctgtcatatacat	360
Db	606	GAGAAAGTAAAGAGCTACATCTATCAGCTAATCAAGGCTATTTCACCTGGTCCATAAGA	665
QY	361	aactgtattcacagagatataaaacctgaaatattctaaactaaagcaaggaataatc	420
Db	666	GATATTGTCATCGATATATAAACACAGAAAATCTCTTAATCAGCCCAATGATGTCTTA	725
QY	421	aagattgtgactctgggtgtgcacaaatcttgcacagagatgccc-----tacacc	474
Db	726	AAACTGTGACATTTGGTTTTCGTAATCTCTCAGAAGGCAATAATGCTAATATACACA	785
QY	475	gattatgtacagagatgtaccagctcctgaactcttctgtggagatcactcagat	534
Db	786	GAGTACGTTGCCACAGATGTATCGGTGCCCAAGCACTTACTTACCTTGGCG---CTCC	842
QY	535	ggttcttcagtcgatatatgggctatgggtgttttttgcagagctcctgcagagcgaag	594
Db	843	GGAAAGTCCGTGGACATGTGTCGGTGGGCTGTATTCTTTGGGGAGCTTAGCGATGGAC	902
QY	595	ccactgtggcctgaaataacagatgtgaccacactttatctgat	638
Db	903	CCTTTATTCTCGAGAAAAGTGAATGTGACCACTTTTACTAT	946
RESULT 10			
LOCUS	AV672492	582 bp	mRNA
DEFINITION	AV672492 Nori Satoh unpublished cDNA library Ciona intestinalis		
ACCESSION	CDNA clone citb5a12 5', mRNA sequence.		
VERSION	AV672492.1 GI:10110491		
KEYWORDS	EST.		
SOURCE	Ciona intestinalis.		
ORGANISM	Ciona intestinalis		
	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;		

REFERENCE 1 (bases 1 to 582)  
 AUTHORS Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.  
 TITLE Expressed genes in *Ciona intestinalis*  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Nori Satoh  
 Department of Zoology  
 Kyoto University  
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan  
 Tel: 81-75-753-4081  
 Fax: 81-75-705-1113  
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES  
 source  
 1..582  
 /organism="Ciona intestinalis"  
 /db\_xref="taxon:7719"  
 /clone\_lib="citb5a12"  
 /clone\_lib="Nori Satoh unpublished cDNA library"  
 /tissue\_type="whole animal"  
 /dev\_stage="tailbud"

BASE COUNT 178 a 100 c 143 g 159 t 2 others  
 ORIGIN

Query Match 16.8%; Score 174.8; DB 10; Length 582;  
 Best Local Similarity 60.3%; Pred. No. 3.6e-29;  
 Matches 323; Conservative 0; Mismatches 209; Indels 4; Gaps 2;

QY 1 atggaaaatagaaaattagtagactggagaagggtctctatggggtgtatcaaa 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 47 ATGGAAATATGACAACCTGGCATGGTGGGTGAAGCGTCCCTATGGGATGTTATGAAA 106  
 QY 61 tgcagaaaacaaacctctggacaagttagctgtgtataaaatttggaaactgaagat 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 107 TGTAAACATAAGGAAGTGGGCAAAATGTGCCATTGAAGTTCTCGAAAGTGAAGAT 166  
 QY 121 gatcctgtgttaagaataatagc-actaagagaataacgtatgttgaagcaattaaaaa 179  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 167 GACAAGATGGTGAAGAAAGATTGGATGCGTGAAGTTGCAAGTTGTGAGAAAACCTCATCA 226  
 QY 180 tccaaactcttggaacctcatcgaggtgttcaggagaaaagaaaatgcattagtttt 239  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 227 CGAAACCTCGTCAACTTAATTGAAGTTTCCGTCGACGTAAAGCTCTTATCTCGCTTT 286  
 QY 240 tgaactactgtatcatacacttttaaatagactggagaagaaacccaatggagtgtga 299  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 287 TGAATTTGTCGACACACACGCTGCTGGATGACCTTGANAATAATCTTAACGGATTAAACGA 346  
 QY 300 tggagtgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatataca 359  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 347 GATGACAGTTCGAAAAATTCCTGGCAAGTCTTGAAGGAGTAGAATTCGTCAACGCCA 406  
 QY 360 taactgtattcagagatatataaaccttgaaaaatattcttaataacaaaggaataat 419  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 407 TATATATACACAGGAGTATCAAAACCTGAAAAATATTTTAAACTCAGCATCGGGGCTCAT 466  
 QY 420 caagatttgacttcgggtttgtcac---aaattctgattccagagagatgcctacacga 476  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 467 CAAACTATGTGACTTTGGTTTGGTCGCGACATTAGCTGCTCTGGTGAGGTTTATACGA 526  
 QY 477 ttatagctacagatggttacccagctctcgaactcttctgtggagatactcagt 532  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 527 TTAATGGCCCACTAGGTGGTATCGTGGCGCCGAGCTGTTGGTCGAGACACCAAAAT 582

RESULT 11  
 A0953706  
 LOCUS A0953706 665 bp DNA GSS 27-JAN-2000  
 DEFINITION Sheared DNA-53L19.TF Sheared DNA Trypanosoma brucei genomic clone  
 ACCSSION A0953706  
 VERSION A0953706.1 GI:6776971  
 KEYWORDS GSS.

SOURCE Trypanosoma brucei.  
 ORGANISM Trypanosoma brucei  
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
 Trypanosoma.  
 REFERENCE 1 (bases 1 to 665)  
 AUTHORS EL-Sayed, N., Zhao, S., Zhao, H., Gill, S., Suh, E., Malek, J., Fujii, C.,  
 Gerrard, C., Leech, V., de Jong, P., Ullu, E., Melville, S., Donelson, J.,  
 Fraser, C. and Adams, M.  
 TITLE Determination of clone end sequences from Trypanosoma brucei GUTat  
 10.1 sheared DNA library  
 JOURNAL Unpublished (1999)  
 COMMENT Other GSSs: Sheared DNA-53L19.TR  
 Contact: Najib M. El-Sayed  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: nelsayed@tigr.org  
 Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared  
 DNA library constructed at TIGR. Clones will be available for  
 distribution through ATCC. Sheared DNA end sequences search page:  
 http://www.tigr.org/tdb/mdb/tbdb/.  
 Seq primer: M13-Forward  
 Class: shotgun.

FEATURES  
 source  
 1..665  
 /organism="Trypanosoma brucei"  
 /strain="TREU927/4 GUTat 10.1"  
 /db\_xref="taxon:5691"  
 /clone\_lib="Sheared DNA-53L19"  
 /clone\_lib="Sheared DNA"  
 /note="Vector: pUC18; Site\_1: SmaI; Constructed at The  
 Institute for Genomic Research (TIGR), Rockville, MD.  
 Genomic DNA isolated from a cloned population of  
 Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically  
 sheared to give a tight size distribution (approx 2 kb).  
 The v + i method used for the library construction is  
 described in detail in Smith, H.O. and Venter, J.C.  
 (Making small insert libraries for whole genome shotgun  
 sequencing projects. In Genome Sequencing: A Practical  
 Approach, eds. M. Vaudin and B. Barell, Oxford University  
 Press, 1999)."

BASE COUNT 183 a 122 c 190 g 170 t  
 ORIGIN

Query Match 16.8%; Score 174.4; DB 13; Length 665;  
 Best Local Similarity 56.4%; Pred. No. 4.4e-29;  
 Matches 325; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

QY 1 atggaaaagtataaaaaattagctaaagactggagaagggtctctatggggtgtattcaaa 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 90 ATGGATCGTACCAAAACACTTGGTATGCTTGGGAGGGGACATTATGTGTAGTTGTGAAG 149  
 QY 61 tgcagaaaacaaacctctggacaagttagctgttbaaaaaatttgygaatctgaagat 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 150 GCGCGGACCGCGGCACATFCACGCATGCTTGCCTATTAATAAGTACAAACAGGCTGAAGAT 209  
 QY 121 gatcctgtttaaagaaaatagcactaagagaaaatcgtatgttgaagcaattaaaaaat 180  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 210 GACGATCACGTCGCCAAAACGTCGCTTCGGGAGGTACGTGTGCTGAACGAGTTACGCCAC 269  
 QY 181 ccaaatcttgtgaacctcatcgagggtgttcaggagagaaaaaagaaatgcattagtttt 240  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 270 CCTAACGTCATAGCCCTGTTAGATGCTATTTCCGCCGCGATGGGAAGCTTTATCTGTTT 329  
 QY 241 gaatactgtgatcatacacttttaaatgagctggaaagaaacccaatgagttgctgat 300  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 330 GAGTATGTTGAGAACACGATTCCTCCAGCTGATTGAGGAAAAGAGATGTGAGATTTCACCC 389  
 QY 301 ggagtgtatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||



Db	390	GATGAGGTACGGCGGTACACCTTTCAGTTATTGACGCGGTGAGTTACTGCCATGCACAC	449			
Qy	361	aactgtattcacagagataaaacactgaataattcttaactaactgaagcaagaataatc	420			
Db	450	AATATCATTTACCGCGGATGTAAGAGCCAGAAAACATCCTAGTATCTAGGAGACGGCGTCTGT	509			
Qy	421	aagatttgtaactctgggtttgcacaaatctcatctccagagatgcctaacacattat	480			
Db	510	AAGTGTGTGCACTTGTGTTCGCCAGGCAAGTGTGTCGTGCGGGGGAATTATACGGAATAT	569			
Qy	481	gtagtacagatgtgtaccagctctctgaactcttctgtgtgggagatactcagtatgtctt	540			
Db	570	CTTGCAACTCGCTGTATCGTCGCCGGAGTTACTAGTTGGTGTATCATATATGCGCAA	629			
Qy	541	tcagtgcataatggctattgttgtttgttgc	576			
Db	630	CGGTGTGTTGGGCTATGTGGCTTTTTC	665			
RESULT 12						
AI385990						
LOCUS	AI385990	495 bp	mRNA			
DEFINITION	m150i06.y1 Stratagene mouse testis (#937308); Mus musculus cDNA					
Clone IMAGE:	515483 5', similar to TR:Q92772 Q92772 P56 KJIANRE					
PROTEIN KINASE.	; mRNA sequence.					
ACCESSION	AI385990					
VERSION	AI385990.1					
KEYWORDS	EST.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Rodentia; Solurognathi; Muridae; Mus; 1 (bases 1 to 495)					
	Marra,M., Hillier,D., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,F., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Willson,R.					
TITLE	The WashU-NCI Mouse EST Project 1999					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Marra M/WashU-NCI Mouse EST project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:309331					
This read is a RESSEQUENCE of a previously sequenced mouse clone						
This read has been verified (found to hit its original self in the correct orientation)						
Possible reversed clone: similarity on wrong strand						
Seq primer: -40RP from Gibco						
High quality sequence stop: 358.						
FEATURES						
source	1..495					
	/organism="Mus musculus"					
	/strain="Inbred CD-1"					
	/db_xref="taxon:10090"					
	/clone_image="IMAGE:515483"					
	/clone_lib="Stratagene mouse testis (#937308)"					
	/sex="males"					
	/tissue_type="testis"					
	/dev_stage="10-12 week old"					
	/lab_host="SOLR (kanamycin resistant)"					
	/note="Organ: testis; Vector: pBluescript SK-; Site:1; EcoRI; Site:2; XhoI; Cloned unidirectionally. Primer: Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "					
BASE COUNT	151	a 100 c 110 g 129 t	5 others			
ORIGIN						

/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: kidney; Vector: PCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library. |"  
BASE COUNT 215 a 143 c 219 g 167 t  
ORIGIN

Query Match 14.9%; Score 154.6; DB 11; Length 744;  
Best Local Similarity 59.2%; Pred. No. 1.2e-24;  
Matches 319; Conservative 0; Mismatches 214; Indels 6; Gaps 3;  
QY 1 atggaaagtataaaatagctaaactggaagggtcttatgggtgtattcaaa 60  
DB 178 ATGGAGATGTATGAACCCCTTGGAAAGTGGGAGGGAAGTTATGACGGGTGATGAA 237  
QY 61 tgcgaagaaacaaacctctggacaagttagtagctgttaaaaaatttgggaactgaagat 120  
DB 238 TCGAAGACATAGGACACTGGCGGAATAGTGGCCATTAGATTTCTATGAGAAACGAGAA 297  
QY 121 gatcctgtgttaagaaatagcactaagagaatacgtatgttggaagcaattaaacat 180  
DB 298 AAATCTGTC---AACAAATTCGAACGAGAGAAATAAAGTTTCTAAAGCAATTCGGCAT 354  
QY 181 ccaactcttgaaacctcatcgaggtgttcaggagaaaaaagaaatgatttagttttt 240  
DB 355 GNAACCTGTCATTTGATTTGAAGCTTTTATGACAAAAGAAAGAAATTCATTTGTATTT 414  
QY 241 gaactctgtatcatcacacttttaaatgagctggaaagaaacccaaatggagttgctgat 300  
DB 415 GAGTTTATGACACACGCTCTTAGATGAGCTGCAGCACTACTGTCACGCACTAGAGAT 474  
QY 301 ggaagtgtcaaaagcgtattatgcaaacactcaagctcttaatttcttgcatacat 360  
DB 475 AAGCGGTGAGAAAGTACTCTTCCAGATCCCTTCGCGCATGAGTACCTGCATATAAT 534  
QY 361 aactgtattcacagagatataaaacctgaaaaatattcttaataactaagcaaggaataac 420  
DB 535 AATATTATTCATCGAGATATAAAGCCTGAGAAATATTTAGTCTCCAGTCAGGAATTACG 594  
QY 421 aagattgtgactcgggttga--caattctgattccagagagagctcctacacgatt 478  
DB 595 AAGCTGTGTGATTTGGGTTTCGAGGACACTAGACGCTCCGAGAGGTTTACACAGACT 654  
QY 479 atgt-agctacagatgtgaccagctcctgaactcttctgtggagatactcaagtgg 536  
DB 655 ACGTGGGCCACACGCTGTGACAGAGCTCCAGAGCTGGTGTGAAAGACACCTCCTACGG 713

RESULT 14  
BG281841  
LOCUS 879 bp mRNA EST 21-FEB-2001  
DEFINITION 602403067F1 NIH\_MGC\_20 Homo sapiens cDNA clone IMAGE:4545400 5',  
mRNA sequence.  
ACCESSION BG281841  
VERSION BG281841.1 GI:13030767  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NH-MGC <http://mgc.nci.nih.gov/>.  
(bases 1 to 879)  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC/DC/DP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LLCMI230 row: e column: 17  
High quality sequence stop: 841.  
Location/Qualifiers  
1. .879  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4545400"  
/clone\_lib="NIH\_MGC\_20"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pORF7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5',  
adaptor: GGCACGAG(C). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 265 a 176 c 197 g 241 t  
ORIGIN

Query Match 14.7%; Score 152.6; DB 11; Length 879;  
Best Local Similarity 54.2%; Pred. No. 3.3e-24;  
Matches 358; Conservative 0; Mismatches 294; Indels 9; Gaps 2;  
QY 1 atggaaagtataaaatagctaaactggaagggtcttatgggtgtattcaaa 60  
DB 102 ATGGAAGATATACCAAAATAGAGAAATTTGAGAGGTACCTATGAGTTGTGTATAAG 161  
QY 61 tgcgaagaaacaaacctctggacaagttagtagctgttaaaaaatttgggaactgaagat 120  
DB 162 GGTAGACACAAACCTCAGGTCAAGTGGTAGCCATGAAAAAATCAGACTGAAAGTGAA 221  
QY 121 gatcctgtgttgaagaaatagcactaagagaatacgtatgttggaagcaattaaacat 180  
DB 222 GAGGAAGGGTTCCTTAGTACTGCAATTCGGGAAATTTCTCTATTAAAGCAACTTCGTCAT 281  
QY 181 ccaactctgtgaaacctcatcgaggtgttcaggagaaaaaagaaatgatttagttttt 240  
DB 282 CCAAAATATAGTCATGCTTCAGGATGTCCTATGAGGATTCACAGGTTCATCTCATCTTT 341  
QY 241 gaactctgtatcatcacacttttaaatgagctggaaagaaacccaaatgg-----agtt 294  
DB 342 GAGTTTCTTCCATGGATCTGAAGAAATACTTGGATTCTATCCCTCTGGTCAGTACATG 401  
QY 295 qctgagggagtgatacaaaacgctattatggcaaacacttcaagctcttaattctgtcat 354  
DB 402 GATTCTTCACTTGTAAAGATTATTTATACCAATCTCAGAGGATGCTGTTTGTTCAC 461  
QY 355 atacataactgtattcacagagatataaaacctgaaaataattcttaactaagcaagga 414  
DB 462 TCTAGAAAGAGTTCTTCACAGAGACTTTAAACCTCAAAATCTCTGATTGATGACAAAGGA 521  
QY 415 ataactaagattgtgacttcgggtttgcacaaattctg---attccaggagatgcctac 471  
DB 522 ACAATTAACTGGCTGATTTTGGCCCTGCCAGACTTTCCGAAATACCTATCAGAGATAT 581  
QY 472 accgattatgtagtcagagatgtgaccagctcctgaactcttctgtgggagatactcag 531  
DB 582 ACATATAGGTAGTAACACTCTGGTACAGATCTCCAGAGATATTGCTGGGTGAGCTCGT 641  
QY 532 tatggttctcagtcgatataatggcctatgggtgtgttttgcagagctcctgcagcggc 591  
DB 642 TACTCAACTCCAGTTGACATTTGGAGTATAGGCACCATATTTCTGGAACCTAGCACTAAG 701  
QY 592 cagcactatggcctggaaaatcagatgtggaccaactttatctgtataatcagaacacta 651  
DB 702 AAACCACTTTCCATGGGATTACAGAAATTTGATCACTCTTCAGGATTTTCAGAGCTTG 761  
QY 652 g 652





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: May 2, 2002, 14:28:48 : Search time 116.84 Seconds  
(without alignments)  
2017.833 Million cell updates/sec

Title: US-09-671-050-9  
Perfect score: 1041  
Sequence: 1 atggaagaatgataaaatt.....aggtaacttcgcgtcaaaagt 1041

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	163.8	15.7	903	2	US-08-874-347-9
2	163.8	15.7	903	3	US-09-093-522-9
3	162	15.6	1476	2	US-08-969-106-1
4	133.2	12.8	1089	1	US-08-154-915-1
5	133.2	12.8	1089	2	US-08-464-517-37
6	133.2	12.8	1089	2	US-08-246-361A-37
7	133.2	12.8	1089	3	US-08-463-772-37
8	133.2	12.8	1089	5	PCT-US93-09945-1
9	128	12.3	1002	1	US-08-463-090B-3
10	116	11.1	1070	1	US-08-463-090B-5
11	103	9.9	2747	2	US-08-874-347-1
12	103	9.9	2747	3	US-09-093-522-1
13	96.2	9.2	1673	4	US-09-347-801-15
14	87.4	8.4	6340	1	US-08-187-785-3
15	87.4	8.4	87350	3	US-08-781-891-79
16	87	8.4	8779	2	US-08-750-703-4
17	86.6	8.3	2624	1	US-08-032-382B-1
18	85.4	8.2	7216	2	US-08-750-703-3
19	84.4	8.1	688	6	5498694-3
20	84.2	8.1	1100	2	US-08-950-449A-18
21	84.2	8.1	1429	4	US-09-347-801-13
22	84.2	8.1	3774	2	US-08-950-449A-11
23	84.2	8.1	3813	1	US-08-469-421-11
24	84.2	8.1	3813	1	US-08-250-975-11
25	84.2	8.1	3813	2	US-08-605-002A-11
26	84.2	8.1	3813	5	PCT-US94-10529-11
27	84.2	8.1	4453	1	US-08-770-761A-4

28	84.2	8.1	4540	1	US-08-770-761A-6
29	84.2	8.1	4621	1	US-08-770-761A-1
30	84.2	8.1	246240	2	US-08-724-394A-20
31	84.2	8.1	246240	2	US-08-724-394A-21
32	84.2	8.1	246240	2	US-08-724-394A-22
33	83.8	8.0	1200	1	US-08-011-398B-3
34	83.8	8.0	1200	1	US-08-464-051-3
35	83.8	8.0	1200	2	US-08-462-498-3
36	83.6	8.0	7720	4	US-09-318-448-5
37	83.4	8.0	35060	3	US-08-814-095-7
38	83.2	8.0	912	2	US-08-627-610-13
39	83.2	8.0	1424	1	US-08-403-634-3
40	83.2	8.0	1424	4	US-08-913-441B-3
41	82.8	8.0	2562	2	US-08-436-771-8
42	82.8	8.0	2562	2	US-08-434-998-8
43	82.8	8.0	2562	2	US-08-487-797-8
44	82.8	8.0	2562	5	PCT-US95-02058-8
45	82.8	8.0	176373	3	US-09-128-155-17

ALIGNMENTS

RESULT 1  
US-08-874-347-9  
; Sequence 9, Application US/08874347  
; Patent No. 5863741  
; GENERAL INFORMATION:  
; APPLICANT: Limper, Andrew H.  
; APPLICANT: Leaf, Edward B.  
; APPLICANT: Thomas, Charles F.  
; APPLICANT: Gustafson, Michael P.  
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
; TITLE OF INVENTION: CARINI  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C., P.A.  
; STREET: 60 South Sixth Street, Suite 3300  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/874,347  
; FILING DATE: 13-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 07039/055001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-335-5070  
; TELEFAX: 612-288-9696  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 903 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 1...900  
; OTHER INFORMATION:  
US-08-874-347-9

Query Match 15.7%; Score 163.8; DB 2; Length 903;  
Best Local Similarity 55.2%; Pred. No. 1.8e-36;  
Matches 365; Conservative 0; Mismatches 287; Indels 9; Gaps 2;

QY 1 atggaagaagtataaaatactagctgaagctggagaggggtcttattgggtgtattcaaa 60  
Db 1 ATGGGCAATATCAGAGGTTAGAGAAAGATTGGAGAGAACTTTATGGAGTTGTTATAAA 60

QY 61 tgcgaagaacaaacctctggacaagtagtctgttaaaaaatttgggaactcgaagat 120  
Db 61 GCAAGGATCTTTGAAGTGGTACAAATTTAGCTCTTAAGAAATCCGATTAGAGCAGAA 120

QY 121 gatcctgtgttaagaataatagcaactaagaatacgtatgttgaagcaattaaaaat 180  
Db 121 GATGAGGGAGTCTCTAGTACAGCAATTCGTGAGATATCATCTTTTGAAGAGATGCACAA 180

QY 181 ccaaatcttggaaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt 240  
Db 181 GATAATGTTGTAAGACTTTTGAATAATAATTCATCAAGAGTCACGTTTATATCTGTTTT 240

QY 241 gaatactgtgatacatatacaacttttaaatgagctggagaagaaccccaaatg-----gagtt 294  
Db 241 GAATTTCTTGATCTTGATTTAAAAAAGTATAGTAATAGTATTCCAAAGGACATGATGCTT 300

QY 295 gctgagtgagtgatcaaaagcgtattatggaacaaacctcgaagctcatttatttctgcat 354  
Db 295 GGTGCAAAATGATTAAAAAGTTTATGTCACAACTTGTATCAGGTGTTAAATATTGTCAT 360

QY 355 atacataactgtattcacagagatataaaacctgaaactgaaatatcttaataaagaagga 414  
Db 355 ATATCGTATCTTCATCGTACGTTGAAACCCACAAAATCTCTTATAGATCGAGAAGGA 420

QY 415 ataatacgaattgtgactcgggtttgtcacaaa---ttctgattccaggagatgcctac 471  
Db 415 AATCTTAAATAGCAGATTTTGGGCTTGCAAGGCGTTTGGTTCCTATCGCTGGTTAT 480

QY 472 accgattatgactacagagatgtaccagagctcctgaagctcttgaagctcgttggagatactcag 531  
Db 472 ACTCATGAGTGTGTACACTTGGTATCGTCTCCAGAGATCTTTTATAGGTTGTCACAA 540

QY 532 tatgttttcagtcgatataatgggctattgtgtgttttgcagagctcctgcagagc 591  
Db 532 TATGCAACAGGCGTTGATATATGAGCATTTGGATGATTTTTCAGAAATGGCTACAAA 600

QY 592 cagccactgtggctgaaatcagatgtgacacacttatctgataatcagaacacta 651  
Db 592 AAGCCATATTTCAGGTGATCTGAAATTCGAAATTCGAAATTTAGAAATATTAGAAATATTA 660

QY 652 g 652  
Db 661 g 661

## RESULT 2

US-09-093-522-9  
; Sequence 9, Application US/09093522  
; Patent No. 6015700  
; GENERAL INFORMATION:  
; APPLICANT: Limper, Andrew H.  
; APPLICANT: Leof, Edward B.  
; APPLICANT: Thomas, Charles F.  
; APPLICANT: Gustafson, Michael P.  
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
; TITLE OF INVENTION: CARINI  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Fish & Richardson P.C., P.A.  
; STREET: 60 South Sixth Street, Suite 3300  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/093,522  
FILING DATE: 08-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/874,347  
FILING DATE: 13-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...900  
OTHER INFORMATION:  
US-09-093-522-9

Query Match 15.7%; Score 163.8; DB 3; Length 903;  
Best Local Similarity 55.2%; Pred. No. 1.8e-36;  
Matches 365; Conservative 0; Mismatches 287; Indels 9; Gaps 2;

QY 1 atggaagaagtataaaatactagctgaagctggagaggggtcttattgggtgtattcaaa 60  
Db 1 ATGGGCAATATCAGAGGTTAGAGAAAGATTGGAGAGAACTTTATGGAGTTGTTATAAA 60

QY 61 tgcgaagaacaaacctctggacaagtagtctgttaaaaaatttgggaactcgaagat 120  
Db 61 GCAAGGATCTTTGAAGTGGTACAAATTTAGCTCTTAAGAAATCCGATTAGAGCAGAA 120

QY 121 gatcctgtgttaagaataatagcaactaagaatacgtatgttgaagcaattaaaaat 180  
Db 121 GATGAGGGAGTCTCTAGTACAGCAATTCGTGAGATATCATCTTTTGAAGAGATGCACAA 180

QY 181 ccaaatcttggaaacctcatcgagggtgttcaggagaaaaaggaaaaatgcatttagtttt 240  
Db 181 GATAATGTTGTAAGACTTTTGAATAATAATTCATCAAGAGTCACGTTTATATCTGTTTT 240

QY 241 gaatactgtgatacatatacaacttttaaatgagctggagaagaaccccaaatg-----gagtt 294  
Db 241 GAATTTCTTGATCTTGATTTAAAAAAGTATAGTAATAGTATTCCAAAGGACATGATGCTT 300

QY 295 gctgagtgagtgatcaaaagcgtattatggaacaaacctcgaagctcatttatttctgcat 354  
Db 295 GGTGCAAAATGATTAAAAAGTTTATGTCACAACTTGTATCAGGTGTTAAATATTGTCAT 360

QY 355 atacataactgtattcacagagatataaaacctgaaactgaaatatcttaataaagaagga 414  
Db 355 TCTCATCGTATCTTCATCGTACGTTGAAACCCACAAAATCTCTTATAGATCGAGAAGGA 420

QY 415 ataatacgaattgtgactcgggtttgtcacaaa---ttctgattccaggagatgcctac 471  
Db 415 AATCTTAAATAGCAGATTTTGGGCTTGCAAGGCGTTTGGTTCCTATCGCTGGTTAT 480

QY 472 accgattatgactacagagatgtaccagagctcctgaagctcttgaagctcgttggagatactcag 531  
Db 472 ACTCATGAGTGTGTACACTTGGTATCGTCTCCAGAGTCTTTTATAGGTTGTCACAA 540

Qy	532	tatggtcttcctcagtcagatataatgggcctatatgtgtgtgttttgcacagcctcctcagcaggc	591
Db	541	TATGCACACAGCGCTTGATATATGAGCAGCTGGATGATATTTTGCAGAAATGGCTACAAAA	600
Qy	592	cagccactgtggcctggaaaaatcagatgtggaccacotttatctgtataatcagaaacacta	651
Db	601	AAGCGATTATTTCCAGSGTGATTCTGAAATTGATGAATATATTTAGATATTAGAAATATTA	660
Qy	652	g 652	
Db	661	G 661	

```

RESULT      3
US-08-969-106-1
; Sequence 1, Application US/08969106
; Patent No. 5986055
; GENERAL INFORMATION:
; APPLICANT: Yang, M.
; APPLICANT: Nandabalan, K.
; APPLICANT: Schulz, V.
; TITLE OF INVENTION: CDK2 INTERACTIONS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,106
; FILING DATE: 13-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: Coding sequence
; LOCATION: 1...894
; OTHER INFORMATION:
; US-08-969-106-1

```

Query Match	15.6%;	Score 162;	DB 2;	Length 1476;
Best Local Similarity	55.0%;	Pred. No. 7e-36;		

Oy	1	atggagaagtagtaaaaaattagctagaactcggagaagggtcttatcggggttgattcctcaaa	60
Db	1	ATGGGAACATTCCAAAGGTGGAAAGATCGGAGAGGCACGTACGGAGTGTGTGTACAAA	60
Oy	61	tgcgagaacaaaacctctggacaagtagtagctgttaaaaaatttctggaatctggaagat	120
Db	61	GCACGAACAACAGTTGACGGGAGGCTGTGGCGCTTTAAGAAATCCGCTCGACACTGAG	120

Qy	121	gatactgttgttaagaaatagcaataagaagaaatacagtattgttgaagcaatataaacacat	181
Dd	121	ACTGAGGGTGTGCCCACTACGTACTGCATCCCGAGAGATCTCTTCGTTAAGGAGCTTAACCAT	180
Qy	181	ccaaatcttgtgaacctcatcgagtggttcaggagaaaagagaaaaatgcatttagttttt	240
Dd	181	CCTAATAATTGTCAAGCTGCTGGATGTCAITTCACACAGAAAAATAAACTCTACTCTGTTT	240
Qy	241	gaatactgtgatcatatcacactttttaatggagctggaa---aagaaaaccacaatggagttgct	297
Dd	241	GAAATTTTGACACCACAAGATCTCAAGAAAATTCATGGATGCTCTGCTCTCACTGGCATTCTCT	300
Qy	298	gatggagtgcatacaaaagcgtattatggcaaacaccticcaagctcttaattttctgcatata	357
Dd	301	CTTCCCCTCATCAAGAGCTATCTCTTCCAGCTGCTCCAGGGCCTAGCTTTCTTCGCCATTCT	360
Qy	358	cataactgtattcacagagatatataaacctgaaaaattctctaataactaagaacaggaaata	417
Dd	361	CATCGGGTCTTCACCCGAGACTTAAACCTTCAGAATCTGCTTTATTAAACAGAGGGGGCC	420
Qy	418	atcaagatttgtgactctgggtttgcacaaattctg---attccaggagatgcctcacacc	474
Dd	421	ATCAAAGCTAGCAGACTTTGGACTAGCCAGCAGAGCTTTTGGAGTCCCTGTTCTGTTACTTACACC	480
Qy	475	gattatgtactacagagatgtaccagagctcctgaactcttgtggagagatactcaagtat	534
Dd	481	CATGAGGTGGTGACCTGTGGTACCGAGCTHCTGGAATCTCTCTGGGCTCGAAATATTAT	540
Qy	535	ggtcttcagtcgatatatgggctattggtgtgttttttcagagctcctcagagggccag	594
Dd	541	TCCACAGCTGTGGACATCTCGAGGCTGGGCTGCATCTTTGTGAGATGTGTACTCGCCGG	600
Qy	595	ccaactgtggctggaataatcagatgtggaccaactttatctcgataatcagaacactag	652
Dd	601	GCCCTGTCCCTGGAGATCTGAGATTGACCACTCTTCCGGATCTTTCGGACTCTGG	658
 RESULT 4			
US-08-154-915-1			
; Sequence 1, Application US/08154915			
; Patent No. 5618669			
; GENERAL INFORMATION:			
; APPLICANT: Beach, David			
; APPLICANT: Xiong, Yue			
; TITLE OF INVENTION: Cyclin Complex Rearrangement and Uses			
; TITLE OF INVENTION: Related Thereto			
; NUMBER OF SEQUENCES: 6			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: LAHIVE & COCKFIELD			
; STREET: 60 State Street			
; CITY: Boston			
; STATE: MA			
; COUNTRY: USA			
; ZIP: 02109			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: ASCII(text)			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/154,915			
; FILING DATE: 19-NOV-1993			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 07/991,997			
; FILING DATE: 17-DEC-1993			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 07/963,308			
; FILING DATE: 16-OCT-1993			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 07/888,178			
; FILING DATE: 26-MAY-1993			
; PRIOR APPLICATION DATA:			

```
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
; US-08-154-915-1
```

```
Query Match      12.8%; Score 133.2; DB 1; Length 1089;
Best Local Similarity 52.3%; Pred. No. 6.5e-28;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

QY 1 atggaaaagtataaaattagctagctggaaggggtcttatggggtgtattcaaa 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ATCGAATAACGAACTGGAAGATTTGGGAAGCCACTACGGAACGTGTTCAAG 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 tgcagaaacaaacctcggacaagtagtctgttaaaaaatttggaaatcgaagat 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCCAAAACCGGAGACTCATGAGATCTGTGGCTCTAAACGGGTGAGCTGATGACGAT 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 121 gatcctgttgaagaaatagcactaagagaatacgtatgttgaagcaataaacaat 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GATGAGGGTGTGCCGAGTTCGCCCTCCGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 181 ccaaatcttggaaacctcatcagggtgttcaggagaaaaggaaaatcatttattttt 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AAGAATCTGACAGGACCTGAAAGACCTGCAAGAGATTTTGCAGAGTTGCAATGGTACCTCGATCCT 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 241 gaatactgtatcacaaccttttaaatgagctggaagaaaccccaaatgagtgatgat 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 ggagtgatacaaaagcgtattatggcaaacacttcaagctcttaatttctgtcatatacat 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAGATTGTAAGTCAATTCCTCTTCCAGCTACTAAAGGGCTGGGATTCGTCTATAGCCGC 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 361 aactgtattcacagagataataaacctgaaaatttcttaataactaagaagaaataatc 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 ATGTGCTACACAGGACCTGAAAGCCCGAGAACTGCTAATAACAGGAAATGGGGAGCTG 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 421 aagatttggacttcgggtttgcacaaa---ttctgattccaggagatgctctacacgat 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AAATTGGCTGATTTTGGCCTGCTCGAGCCTTTGGGATTCCTCCGCTGTTACTCAGCT 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 478 tatgtagctacagatgtgtaccagagctcctgaactcttctgtggagatactcagtagt 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 GAGGTGGTCACACTGTGGTACCGCCCGCTGCTATCTTTGGGGCCAAAGCTGTACTCC 552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 538 tctcagtcgatataatgggtatttgggtgttttgcagagct---cctgcagagccag 594
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 ACCTCATCGACATGTGGTCAGCGCTGCTATCTTTGGGGCCAAAGCTGTACTCC 612
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 595 ccactgtggcctgaaaatacagatgtggacaactttactgataataagaacactag 652
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 CCTCTTTTCCCGCAATGATGTCATGACCAAGCTTGAGAGGATCTTCCGACTGCTGC 670
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 5

US-08-464-517-37

```
; Sequence 37, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
; US-08-464-517-37
```

```
Query Match      12.8%; Score 133.2; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 6.5e-28;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;
```

```
QY 1 atggaaaagtataaaattagctagctggaaggggtcttatggggtgtattcaaa 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ATCGAATAACGAACTGGAAGATTTGGGAAGCCACTACGGAACGTGTTCAAG 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 tgcagaaacaaacctcggacaagtagtctgttaaaaaatttggaaatcgaagat 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCCAAAACCGGAGACTCATGAGATCTGTGGCTCTAAACGGGTGAGCTGATGACGAT 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 121 gatcctgttgaagaaatagcactaagagaatacgtatgttgaagcaataaacaat 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GATGAGGGTGTGCCGAGTTCGCCCTCCGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 181 ccaaatcttggaaacctcagaggtgttccaggagaaaaggaaaatgcatatttttt 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AAGAATCTGCTAGGCTTCATGAGTCTCTGCACAGCGCAAGAGCTGACTTTGGTTTTT 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 241 gaatactgtatcacaaccttttaaatgagctggaagaaaccccaaatgagtgatgat 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GAATCTGTGACCAAGACCTGAAGAAGTATTTTGACAGTTTGAATGTGACCTGACCTGCT 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```
QY 301 ggagtgatcaaaagcgattatggtgcaaacacttcaagctcttaatttctgtcatatacat 360
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAGATTGTAAAGTCATTCCTCTCCAGCTACTAAAGGGCTGGGATCTGTCTATAGCCGC 372
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 aactgtattcacagagatataaaacctgaaatatcttaataactaaagaaggaataatc 420
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 AATGTGCTACAGGGACCTGAAGCCAGAACCTGTCTAATAACAGGAATGGGGAGCTG 432
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 aagattgtgacttcgggttgcacaaa---ttctgattccaggagatgcctcacacgat 477
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AATGGCTGATTTGGCCCTGGCTCGAGCCITTTGGGATTTCCGCTCGCTGTACTAGCT 492
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 tatgtagctacagagatggtaccagagctcctgaacttcttggggagatactcaagtatggt 537
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 GAGGTGGTCACATGTGTACCGCCACCGGATGTCCTCTTTGGGGCCCAAGCTGTACTCC 552
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 538 tcttcagtgatataatgggtattgttgggttgggttgggttgggttgggttgggttgggt 594
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 AGTCCATGCACATGTGTACCGCCACCGGCTGCATCTTTGCAGAGCTGGCCAAATGCTGGCGG 612
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 ccactgtgacctgaaatacagatgtggaccaactttatctgataatcagaacactag 652
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 CCTCTTTTCCCGCAATGATGTCGATGACCAAGTTGAAGAGGATCTTCGGACTGCTGG 670
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 6
US-08-246-361A-37
; Sequence 37, Application US/08246361A
; Patent No. 598582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
US-08-246-361A-37

Query Match 12.8%; Score 133.2; DB 2; Length 1089;
Best Local Similarity 52.3%; Pred. No. 6.5e-28;
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

QY 1 atggaagaatgaaaaattagctaaagctgagaagggtcttattgggttgattatcaaa 60
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 ATCCAGAAATACGAGAAACTGGAAGAAGATTTGGGAAGGCACCTACGGAACCTGTTCAAG 72
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 tgagaaacacaaacctctggacaagtagtagctgttaaaaaatttggtaactcgaagat 120
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCCAAACCGGGAGACTCATGAGATCGTGGCTCTAAACCGGGTGAGGCTGGTGGAGAT 132
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 gatcctgtgttaagaaatagcactaaagaaatacgtatgttgaagcaattaaaaacat 180
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 GATGAGGTGTCCGAGTTCCGCCCTCCGGGAGATCTGCCTACTCAAGGAGCTGAAGCAC 192
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 ccaaatcttggtaacctcatcaggtgttcaggagaaaaagaaaaatcatttagtttt 240
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AAGAACATCGTCAGGCTTCATGACGTCCTGCACAGCCGACAGGAAGCTGACTTTGGTTTT 252
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 gaatactgtgatacacacttttaaatgagctgtgaaagaaaccccaatggagttgctgat 300
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GAATTTCTGTGACCAGGACCTGAAGAAGATATTTTGACAGTTTGCAATGGTGACCTCGATCCT 312
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 ggagtgatcaaaagcgattatggtgcaaacacttcaagctcttaatttctgtcatatacat 360
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAGATTGTAAAGTCATTCCTCTCCAGCTACTAAAGGGCTGGGATCTGTCTATAGCCGC 372
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 aactgtattcacagagatataaaacctgaaatatcttaataactaaagaaggaataatc 420
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 AATGTGCTACAGGGACCTGAAGCCAGAACCTGCTCTAATAACAGGAATGGGGAGCTG 432
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 aagattgtgacttcgggttgcacaaa---ttctgattccaggagatgcctcacacgat 477
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AATGGCTGATTTGGCCCTGGCTCGAGCCITTTGGGATTTCCGCTCGCTGTACTAGCT 492
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 tatgtagctacagagatggtaccagagctcctgaacttcttggggagatactcagtatggt 537
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 GAGGTGGTCACATGTGTACCGCCACCGGATGTCCTCTTTGGGGCCCAAGCTGTACTCC 552
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 538 tcttcagtgatataatgggtattgttgggttgggttgggttgggttgggttgggttgggt 594
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 AGTCCATGCACATGTGTACCGCCACCGGCTGCATCTTTGCAGAGCTGGCCAAATGCTGGCGG 612
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 ccactgtgacctgaaatacagatgtggaccaactttatctgataatcagaacactag 652
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 CCTCTTTTCCCGCAATGATGTCGATGACCAAGTTGAAGAGGATCTTCGGACTGCTGG 670
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 7
US-08-463-772-37
; Sequence 37, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
US-08-463-772-37
```

Query Match 12.8%; Score 133.2; DB 3; Length 1089;  
Best Local Similarity 52.3%; Pred. No. 6.5e-28;  
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

```

QY 1 atggaagaagtataaaattagcactgagaagaagggtcttatggggtgtattcaaa 60
DB 13 ATCGAAGATACGAGAACTGGGAAAGATTGGGGAAGGACCTACGGAACCTGTGTTCAAG 72
QY 61 tgcagaacaaacctcggacaagtagtagctgttaaaaaatttgtgaatctgaagat 120
DB 73 GCCAAACCCGGAGACTCATGAGATCGTGGCTCTAAACCGGTGAGCTGGATGACGAT 132
QY 121 gatcctgtgttaagaaatagcactaagagaatacgtatgttgaagcaatataaacat 180
DB 133 GATGAGGGTGTGCCGCTTCATGACGCTCTGACAGGACGACAAAGAGCTGAAAGCAC 192
QY 181 ccaaatcttgtgaacctcattatgagcaaacctcgaagctcttaattctgtgcataacat 240
DB 193 AAGAATGTAAAGTCATTCTCTCCAGCTACTAAAGGGCTGGGATCTGTCTATAGCCGC 252
QY 241 gaatactgtgatacacatttaaatgagctggaagaaacccaataagagtgctgat 300
DB 253 GAATCTGTGACAGGACCTGGAAGAGATTTTTCACAGTTGCAATGTGACCTCGATCCT 312
QY 301 ggagtgatcaaacgctattatggcaaacacttcaagctcttaattctgtgcataacat 360
DB 313 GAGATGTAAAGTCATTCTCTCCAGCTACTAAAGGGCTGGGATCTGTCTATAGCCGC 372
QY 361 aactgtattcacagagataaaaacctgaaattattctaataactaagcagaataatc 420
DB 373 ATGTGCTACACAGGACCTGAAGCCCGAGACCTGCTAATAACAGGAATGGGAGCTG 432
QY 421 agagatttgtgactcgggtttgcacaaa---ttctgattccagagagatgctcacacgat 477
DB 433 AAATGTGGTGAATTTGGCTCGAGCCTTTGGGATTTCCCGCTGCTTACTCAGCT 492
QY 478 tatgtagctacgagatggtaccgagctcctgaactcttgtggagatactcagtatagt 537
DB 493 GAGGTGGTCACACTGTGGTACCGCCCGGATGTCCTCTTTGGGGCCCAAGCTGTACTCC 552
```

```

QY 538 tcttcagtcgatatatatgggctatattggtgtgtttttgcagagact---cctgacagggccag 594
DB 553 ACGTCCATCGACATGTGGTCACCGGCTGCATCTTTGCAGAGCTGGCAATGCTGGCGCG 612
QY 595 ccactgtgacctggaaaaatcagatgtgaccaactttatctgataatacagaacactag 652
DB 613 CCTCTTTTCCCGGCAATGATGTCGATGACCCAGTTGAAGAGGATCTTCCGACTGCTGG 670
```

```

RESULT 8
PCT-US93-09945-1
; Sequence 1, Application PC/TUS9309945
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Cyclin Complex Rearrangement and Uses Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/09945
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,997
; FILING DATE: 17-DEC-1992
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..888
PCT-US93-09945-1
```

Query Match 12.8%; Score 133.2; DB 5; Length 1089;  
Best Local Similarity 52.3%; Pred. No. 6.5e-28;  
Matches 344; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

```

QY 1 atggaagaagtataaaattagcactgagaagaagggtcttatggggtgtattcaaa 60
DB 13 ATCGAAGATACGAGAACTGGGAAAGATTGGGGAAGGACCTACGGAACCTGTGTTCAAG 72
QY 61 tgcagaacaaacctcggacaagtagtagctgttataaaaaatttgtggaatctgaagat 120
DB 73 GCCAAACCCGGAGACTCATGAGATCGTGGCTCTAAACCGGTGAGCTGGATGACGAT 132
QY 121 gatcctgtgttaagaaatagcactaagagaatacgtatgttgaagcaatataaacat 180
DB 133 GATGAGGGTGTGCCGCTTCATGACGCTCTGACAGGACGACAAAGAGCTGAAAGCAC 192
QY 181 ccaaatcttgtgaacctcattatgagcaaaccttcaagctcttaattctgtgcataacat 240
DB 193 AAGAATGTAAAGTCATTCTCTCCAGCTACTAAAGGGCTGGGATCTGTCTATAGCCGC 252
QY 241 gaatactgtgatacacatttttaaatgagctggaagaaacccaataagagtgctgat 300
DB 253 GAATCTGTGACAGGACCTGAAGAGATTTTTCACAGTTGCAATGTGACCTCGATCCT 312
QY 301 ggagtgatcaaacgctattatggcaaacacttcaagctcttaattctgtgcataacat 360
DB 313 GAGATGTAAAGTCATTCTCTCCAGCTACTAAAGGGCTGGGATCTGTCTATAGCCGC 372
```

QY 361 aactgtattacagagatataaaacactgaaaaataatcttaataactaagcaaggaataatc 420  
 Db 373 AATGTGCTACAGAGGAGCTGAAGCCCCAGAACCTGTCTAATAAACAGGAATGGGAGCTG 432  
 QY 421 aagatttgacttcgggtttgcacaaa---ctctgattccagagagatgcctacaccgat 477  
 Db 433 AAATGGCTGATTTTGGCTGGCTCGAGCCCTTTGGGATTCCTCTTTGGGGCCAACTGTACTCAGCT 492  
 QY 478 tatgtagctacagatggtaccagctcctgaaactcttctgtggagatactcaagtatggt 537  
 Db 493 GAGTGGTGCACACTGTGTACCGCCACCGGATGTCTCTTTGGGGCCAACTGTACTCAGCT 552  
 QY 538 tcttcagctgatataggtgattggtgtgttttgcagagct---cctgacagggccag 594  
 Db 553 ACGTCCATCGACATGTGTGAGCCGGCTGCATCTTTGACAGAGCTGGCCAACTGTGGGGG 612  
 QY 595 ccactggtgctggaataatcagatgtggacaaactttatctgataatcagaacactag 652  
 Db 613 CCTCTTTTCCCGCAATGATGTGATGACCGAGTTGAAGAGGATCTTCGAGCTGTGG 670

RESULT 9  
 US-08-463-090B-3  
 ; Sequence 3, Application US/08463090B  
 ; Patent No. 5801015  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cottarel, Guillaume  
 ; APPLICANT: Damagnez, Veronique  
 ; APPLICANT: Draetta, Guilio  
 ; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins from  
 ; HUMAN PATHOGENS, AND USES RELATED THERETO  
 ; NUMBER OF SEQUENCES: 25  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley, Hoag & Eliot, LLP  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII (text)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/463,090B  
 FILING DATE: 05-JUN-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Vincent, Matthew P.  
 REGISTRATION NUMBER: 36,709  
 REFERENCE/DOCKET NUMBER: MIV032.01  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 832-1299  
 TELEFAX: (617) 832-7000  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1002 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 43..993

US-08-463-090B-3  
 Query Match 12.3%; Score 128; DB 1; Length 1002;  
 Best Local Similarity 53.9%; Pred. No. 1.8e-26;  
 Matches 357; Conservative 0; Mismatches 290; Indels 15; Gaps 4;  
 QY 8 agtatgaaaataagtaagactggagaggggtcttatggtgtatttcaa---atgca 64

Db 59 ATTATCAAGCTCAGAAAAAGTCGAGAAAGTACTTATGGGGTGTGTTTATAAAGCATTAG 118  
 QY 65 gaacaaaaacctctcggaagtagctgttataaaatttgggaactctgaagatgac 124  
 Db 119 ATACCAAGCACAAATAAGAGTTTGGCATTAAGAAAAATTCGATTAGAAACAGAAATG 178  
 QY 125 ctgttctgaagaaatgcgactaagagaaatacgtatgttgaagcaatataaacaatccaa 194  
 Db 179 AAGGTGTACTAGTAGACCGCATTTAGAAAAATCTCGTTATTAAAGAAATGAAAGATGATA 238  
 QY 185 atcttgaacactcatcaggtgttcaggagaaaa---aaggaaaaatgcatttagttttg 241  
 Db 239 ATATCGTTCGATTATATGATATTATTCATTTCAGATTCTCATAAATATATATTATTAGT 298  
 QY 242 aactgtgtatcatcaccttcttaaatgagctggaaagaaacccaaatggagttgctgatg 301  
 Db 299 AATTTTGGATTATGATTTAAAGAAATATATGAAAGTATTCCTCAAGGAGTTGGACTAG 358  
 QY 302 gag-----tgatcaaaagcgtattatggcaaacacttcaagctcttaatttctgtcata 355  
 Db 359 GGCTAATATGATAAAGAGATTATGAATCAATTAATTCGAGGTATTAACACATTGTCTATT 418  
 QY 356 tacataactgtattcacagagatataaaacctgaaataatcttaataactaagcaagaaa 415  
 Db 419 CTCATCGAGTTTACATCGTATTAACACCAACAAAAATTTATTGATTGATAAAGAGGGA 478  
 QY 416 taatcaagatttgcactcgggtttgcacaaa---ttctgattccagagatgcctaca 472  
 Db 479 ATTTAAAAATTACAGATTTTGGATTAGCTCGAGCATTTGGAGTTCCATTAGAGCATATA 538  
 QY 473 ccgattatagctacagagatggtaccagagctcctgaactctcttggagagatactcagt 532  
 Db 539 CTCATGAAGTTTGTCTACTTTATGTTATCGAGCTCCGAAATCTTTGTTAGGAGGGAACAAT 598  
 QY 533 atggttctcagtcgatatatgggtctgtgtgtgttttgcagagctcctgcacaggcc 592  
 Db 599 ATTCCACTGGGTAGATATGTTGGTCTGTTGGATGTATATTGCTGAAATGTGTAATAGGA 658  
 QY 593 agccactgtgctggaataatcagatgtggaccacactttatctgataatcagacacactag 652  
 Db 659 AACCATTATTTCCCTGGTGTATTCAGAAATTTGATGAAATTTTCCGAATTTTCCGAATTTTAG 718  
 QY 653 ta 654  
 Db 719 GA 720

RESULT 10  
 US-08-463-090B-5  
 ; Sequence 5, Application US/08463090B  
 ; Patent No. 5801015  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cottarel, Guillaume  
 ; APPLICANT: Damagnez, Veronique  
 ; APPLICANT: Draetta, Guilio  
 ; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins from  
 ; HUMAN PATHOGENS, AND USES RELATED THERETO  
 ; NUMBER OF SEQUENCES: 25  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley, Hoag & Eliot, LLP  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII (text)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/463,090B



```
QY 353 atatacataactgtattcacagagatataaaacctgaaataattcttaactaactaagaag 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1411 ATTCTATCGTATTCTTCATCGTGACTTGAAGCCCAAAATCTTCTATTGATCGGAGAAG 1470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 gaataatcaagattgtgacttcgggtttgcac---aaattctgatlccagagagatgcct 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1471 GAAATCTTAAACTTGGCTGATTTCCGGCTTGCTCGGCATTTGGTGTTCCTTTACGTGGTT 1530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 470 acaccgattatgtacacagatagttaccagactcctgaactctcttctgtggagatactc 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1531 ATACTCATGAGTTGTACACTTGGTATCGTGTCCGAAGTCTTTTAGTGGTGCAC 1590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 530 agtatggtcttcagtcgatatatatgggctatgtgtgtgtttttgcagagactcctgaacag 589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1591 AATATGCAACAGCACTTGATATTGGAGCATTTGGATGTATTTTTGCTGAAATGGCTACAA 1650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 590 gccagccactgtggcctgaaatacagatgtggaccaactttatctgataatacagaaca 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1651 AAAAACCGTTATTCCAGCGGATTCTGAAATTGATGAAATATTCAGAAATATTAGTGCA 1709
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 12
US-09-093-522-1
; Sequence 1, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; TITLE OF INVENTION: CARINII
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/093,522
; FILING DATE: 08-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,347
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 857...893
; OTHER INFORMATION: (A
; LOCATION: 942...1096
```

```
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1147...1194
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1242...1254
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1311...1706
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1751...1891
; OTHER INFORMATION:
; NAME/KEY: Coding Sequence
; LOCATION: 1944...2048
; OTHER INFORMATION:
; US-09-093-522-1

Query Match 9.9%; Score 103; DB 3; Length 2747;
Best Local Similarity 57.4%; Pred. No. 2.6e-19;
Matches 206; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

QY 293 ttgctgatggagtgatcaaaaagcgtattatgagcaaacactcaagctcttaattctgtc 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1351 TTGGAGCTGAAATGATCAAAAAATTTATGTCTCAACTTGTATCAGGGTTTAAATATTGTCT 1410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 353 atatacataactgtattcacagagatataaaacctgaaataattcttaactaactaagaag 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1411 ATTCTCATCGTATTCTTCATCGTGACTTGAAGCCCAAAATCTTCTATTGATCGAGAG 1470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 gaataatcaagattgtgacttcgggtttgcac---aaattctgatlccagagagatgcct 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1471 GAAATCTTAAACTTGGCTGATTTCCGGCTTGCTCGGCATTTGGTGTTCCTTTACGTGGTT 1530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 470 acaccgattatgtacacagatgtgaccagactcctgaactctctgtggagatactc 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1531 ATACTCATGAGGTGTGTACACTTTGGTATCGTGTCCAGAGTCTTTTAGTGGTGCAC 1590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 530 agtatggtcttcagtcgatatatatgggctatgtgtgtgtttttgcagagactcctgaacag 589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1591 AATATGCAACAGCACTTGATATTGGAGCATTTGGATGTATTTTTGCTGAAATGGCTACAA 1650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 590 gccagccactgtggcctgaaatacagatgtggaccaactttatctgataatacagaaca 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1651 AAAAACCGTTATTTCAGCGGATTCTGAAATTCATGAAATATTCAGAAATATTAGTGCA 1709
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-347-801-15
; Sequence 15, Application US/09347801
; Patent No. 6262345
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; TITLE OF INVENTION: Plant Protein Kinases
; FILE REFERENCE: BH-1171
; CURRENT APPLICATION NUMBER: US/09/347,801
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,438
; EARLIER FILING DATE: July 10, 1998
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1349)
; US-09-347-801-15
```

```

Query Match          9.2%; Score 96.2; DB 4; Length 1673;
Best Local Similarity 61.7%; Pred. No. 1.6e-17;
Matches 171; Conservative 0; Mismatches 103; Indels 3; Gaps 1;

QY 370 cacagacatataaacctgaaataattcttaataaactaaagcaa---ggaataatcaagatt 426
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 599 cacaggagtgtaaacaccacaaaattgttggtagctctctaaacctcaagtcacagatc 658
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 427 tgtgaacttcgggtttgcacaaattcttgattccagagagatgctctacacagcattatgtagct 486
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 659 tgtgaacttcgggaagtgcacaaagtcttggtacacctgtgtaaccccaacatagacatacatatgc 718
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 487 acgagatggtaccagagctcctggaactcttctgtggagagatactcagtatggtcttcacagtc 546
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 719 tctgcgtactactgcgtcctcgagctcatatttggcgactgaatatacaacttcaata 778
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 547 gatataatgggttatgtgtgtgtttttgagagagctcctgcagagcagccactgtggcct 606
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 779 gacataatggtcagctggatgtgttcttgacagagctactcttctgtgcagcctctgtttcca 838
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 607 ggaataatcagatgtggaccacttctatctcgataatca 643
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 839 ggagagactcgggtgatcagctagtgagattatca 875
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-08-187-785-3/c
; Sequence 3, Application US/08187785
; Patent No. 5756476
; GENERAL INFORMATION:
; APPLICANT: Epstein, Stephen
; APPLICANT: Unger, Ellis
; APPLICANT: Speir, Edith
; TITLE OF INVENTION: Inhibition of No. 5756476-Transformed Cell
; TITLE OF INVENTION: Proliferation Using Anti-Sense Oligonucleotides
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson, and Bear
; STREET: 620 Newport Center Dr. Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,785
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/821,415
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altmap, Daniel E.
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: NIH001.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6340 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human

; IMMEDIATE SOURCE:
; CLONE: PCNA Genomic Clone
US-08-187-785-3

Query Match          8.4%; Score 87.4; DB 1; Length 6340;
Best Local Similarity 75.2%; Pred. No. 8.7e-15;
Matches 109; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 650 taatagacagcgggtttcgccatgttgaccagggctggtctcgaactcttgacgtcaagtg 709
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 TAGTAGAGACGGGGTTTCACCATGTGTGGCCAGGCTGTGTGAACACCTGACCTCAGGTG 167
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 710 atccacctgcgtagcctctcaaaagtgtggaattacaggaataataatcccaagacatc 769
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 ATCCACCTGCCTCGSCCTCTCAAAGTGTGTGGGATTACAGGCTGTGAGCCACCAACGCCACG 107
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 770 aatcaatctttaaagaataacgggtt 794
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 106 CAGAAATTTTCAATTGTTACAGGAT 82
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-08-781-891-79
; Sequence 79, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-781-891-79

Query Match          8.4%; Score 87.4; DB 3; Length 87350;
Best Local Similarity 89.5%; Pred. No. 3e-14;
Matches 94; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 648 actagtagacagcgggtttcgccatgttgaccagcgtgtctcgaaactcttgacgtcaag 707
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 85777 ATTAGTAGACAGCGGATTGCGCATGTTGATCAGGCTGGTCTCGAAGCTCCTGACCTCAAG 85836  
Qy 708 tgatccactgcgtagcctctcaagtgtcggaattacagaaa 752  
|||||  
Db 85837 TAATCCACCTGCCTTAGCTTCCCAAGTGTGGGATTACATAGA 85881

Search completed: May 2, 2002, 17:09:45  
Job time: 9657 sec

